

The Structure of English Tonal Morphemes

A thesis

submitted in partial fulfilment
of the requirements for the degree

of

Doctor of Philosophy

in

the University of Canterbury

by

Douglas Cairns Haggo

University of Canterbury

1987

PE
1139.5
.H145
1987

Contents

Abstract	vi
Acknowledgements	vii
 Introduction	 1
 1. Tonal Morphemes	 8
1. Autosegmental phonology	9
1.1. Background	9
1.2. Representations	11
1.3. Rules	12
2. Prominence	15
2.1. Three kinds	15
2.2. Prominence and phrasing	18
3. English tonal morphemes	19
3.1. The Drop	20
3.2. The Low Bounce	26
3.3. The Dipper.....	28
4. The term <i>tonal morpheme</i>	31
4.1. Tonal	32
4.1.1. Phonology of tone	32
4.1.2. Tone and pitch	33
4.2. Morpheme	38
4.2.1. Grammatical unit	38
4.2.2. Minimum grammatical unit	40
4.2.3. Stylistic morphemes	43
 2. Variable Stress	 46
1. The Problem	46
2. An Explanation	50

2.1. Lexical prominence features	51
2.2. Nuclear Accentuation Rule	52
2.3. Prenuclear Accentuation Rule	55
2.4. Application to the problem	57
3. Other explanations	60
3.1. Gimson	61
3.2. Vanderslice and Ladefoged	63
3.3. Liberman and Prince	67
4. Revisions	71
4.1. A new NAR	72
4.2. Redundant features	75
3. Obligatory Prenuclear Tones	79
1. The Low Bounce	79
2. The Long Jump	84
3. The Switchback	85
3.1. Switchback and Dipper	86
3.2. Competing hypotheses	88
3.3. Tone association rules	93
4. The Fifth Tone	98
1. Need for five tones	98
2. Trager and Smith	101
2.1. Comparison	101
2.2. Terminal junctures	104
2.3. Redundant and missing pitches	108
3. Other four-tone analyses	110
3.1. Pike	110
3.2. Wells	111
5. Postnuclear Accent Tones	113
1. Tone spreading solution	115
1.1. How it works	115

1.2. Weaknesses	116
2. Accent tone solution	120
2.1. First alternative	121
2.2. Second alternative	122
3. Tune-sensitive accentuation rules	125
3.1. Introduction	125
3.2. Stock auction presale tune	127
3.3. Conclusion	130
 6. Levels versus Contours	 132
1. Related hypotheses	133
1.1. Phonological segment	134
1.2. Phonological unit	135
1.3. Nuclear tones as units	136
1.4. Tunes as units	137
2. Arguments against contours	138
3. Arguments for atomic contours	141
4. Contourist critiques of levels	143
4.1. Bolinger/Crystal	143
4.2. Cruttenden	145
4.3. Ladd	149
4.3.1. Stylized intonation	149
4.3.2. Pitch direction generalizations	155
 7. Binuclear Tunes	 159
1. Three proposals	161
1.1. O'Connor and Arnold	161
1.2. Halliday	163
1.3. Crystal	164
1.3.1. Fall-plus-rise	164
1.3.2. Rise-plus-fall	167
2. Reasons for binuclear tunes	172
2.1. No boundary	172

2.1.1. Pitch change criterion	173
2.1.2. Pause criterion	175
2.1.3. Phonetic criteria	176
3. Fall-plus-rise as two morphemes	181
3.1. Second nucleus tone	182
3.2. Boundary between the nucleus tones	184
8. Anuclear Tunes	188
1. Status of initial tone	190
2. Status of 1-tone	192
3. Location of accent tone	195
4. Absence of nucleus tone	197
5. Optional tones	199
6. Other uses of the contour	201
References	206

Abstract

Several attractively simple hypotheses about the structure of English tonal morphemes, or intonational tunes, are investigated. In some cases I argue that the hypothesis can be maintained; in others I conclude that it must be abandoned. In the course of the investigation I make proposals both about the representation of particular tonal morphemes and about the rules which associate their constituent tones with the syllables of phonological phrases. The operation of these tone association rules is demonstrated in Chapter 1, which also includes a defence of the hypothesis that the units in question are morphemes. In Chapter 2 I support the analysis of prominence required for the application of the tone association rules by showing that it offers a natural account of the phenomenon of stress shift.

A tune is identified in the third chapter which falsifies the hypothesis that the nucleus tone is the first obligatory tone in a tonal morpheme. In Chapter 4 I attack the hypothesis that four tones suffice to model the tunes of English, showing that five tones are necessary. I argue in Chapter 5 that the "calling tune" requires us to abandon the hypothesis that there are no accent tones following the nucleus tone of a tonal morpheme. In the sixth chapter I defend the hypothesis that tonal morphemes consist entirely of level tones.

In the last two chapters I take up two complementary kinds of apparent exception to the hypothesis that each tonal morpheme contains one and only one nucleus tone. I discuss and dismiss in Chapter 7 claims that some tonal morphemes must contain more than one nucleus tone. In Chapter 8 I conclude that the opposite type of exception-- a tonal morpheme with no nucleus tone-- does exist, in the form of the "contradiction contour", which is also exceptional in other ways.

Acknowledgements

This thesis could not have been written without the long-term support of Derek Davy, Kon Kuiper, and Regina Haggio. I am deeply grateful to all three. Thanks are also due to Andrew Carstairs for his helpful comments on an earlier version.

Introduction

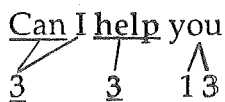
Several attractively simple hypotheses about the structure of English tonal morphemes, or intonational tunes, are investigated in the following study. In some cases I argue that the generalization can be maintained; in others I conclude that it must be abandoned. In the course of evaluating these hypotheses I make and compare proposals about the form of particular tunes, using phonetic data in the shape of laryngograph displays to illustrate and support my claims. That is, I will be trying to determine the best representation for particular morphemes. I will also have proposals to make about the character of the rules which link the constituent tones of these morphemes to the syllables of the associated phrases.

The study of intonation is traditionally described as part of phonology, e.g. suprasegmental phonology. Since both tones and syllables are phonological units, proposals about the tone association rules which link them do appear to belong to phonology. Investigation of the form of tonal morphemes, however, would appear to be more appropriately described as morphology. Students of so-called segmental phonology and morphology have found a great deal of give and take between the two. That is, decisions about the best representation of particular morphemes depend on assumptions about the nature of phonological rules, and decisions about the correct formulation of phonological rules depend on assumptions about the form of particular morphemes. The same is true with tonal morphemes and tone association rules. Attempting to find the best representation for a tonal morpheme may lead to revision of the tone association rules. This two-way traffic is typical of research not only in linguistics but in other fields as well. As Allen (1973:16) puts it, "When one seeks to increase one's understanding of *x* by reference to the facts of an already known *y*, the process not uncommonly suggests some re-interpretation of the latter".

The questions under investigation are framed in terms of a theory which represents the tune of each phonological phrase as one of a limited number of tonal morphemes. The analysis is fairly traditional, but couched in the relatively new framework of autosegmental phonology. Before proceeding to summarize the contents of this study chapter by chapter, I will informally outline the main features of the theory presupposed herein. The summary will be followed by some information about the records of the pitch of utterances which illustrate this work.

In the interest of readability I will omit locutions such as "I hypothesize that" from the following sketch. The syllables of spoken English are organized into larger phonological units which I call phonological phrases. At the ends of stretches of speech which instantiate phonological phrases, native speakers of the language perceive a pause. Every well-formed phonological phrase in an English utterance carries a tune from an inventory of English tunes, or tonal morphemes. The number of tunes in each dialect is fairly small-- closer to ten than to a hundred.

Each tonal morpheme consists of a sequence of notes, or tones as I call them. Each tone is represented by a number. The higher the number, the higher the pitch of a syllable linked to that tone, other things being equal. There are five tones, but tones 1 and 3 appear more often than the others in English tunes. The tune which I call the Dipper, for example, contains only these two tones: (1)(3)313. Parentheses indicate tones which are optional, in the sense that they do not occur if the first syllable of the associated phrase is the most prominent in the phrase. If this tune is used with an utterance of *Can I help you?* in which the first syllable is prominent, but less so than the third, the first optional tone (1) is absent, but the second one (3) is present:



Tone association lines connect the tones of a tonal morpheme with the syllables of a phonological phrase. Every syllable in a phonological phrase must be attached to a tone, and every obligatory tone must be attached to a syllable. As the diagram shows, a single tone can be attached to more than one

syllable: the first 3-tone is attached to both *can* and *I*. Conversely, one syllable can be attached to more than one tone. This too, is illustrated above: *you* is attached to both a 1-tone and a 3-tone, in that order. This means that its pitch is low at the beginning and relatively high at the end.

The tone association lines are drawn, as it were, by tone association rules, of which there are two kinds. Marked tones are attached by the first kind, known as initial tone association rules. Some of these require information about the prominence of syllables: the distribution of tones across the syllables of a phrase depends largely on the prominence of the syllables. This information is represented in the transcription above by underlining and boldface. For example, the nucleus tone of a tonal morpheme, represented in boldface, gets attached to the nuclear syllable, i.e. the most prominent syllable of a phonological phrase, which is marked in the same way. An accent tone, which is underlined in plaintext, gets attached to a syllable which is accented but not nuclear.

The tune described above contains one accent tone and one nucleus tone, the former preceding the latter. We will adopt the working hypothesis that the same is true of all English tonal morphemes. The Dipper is also typical in that the nucleus tone is the first of the obligatory tones. It also contains an outlined tone, called an end tone, which gets attached to the last syllable in the phrase. Not all tonal morphemes have such a tone. After the marked tones of this tune have been attached to the sample phrase, the the two associated sequences will be represented as follows.

Can I **help** you
 3 3 1 3

We still have some unattached syllables and tones here. Unmarked tones are attached by tone association rules of the second kind, known as well-formedness rules, which operate in a well-defined manner to make every representation well formed. That is, they ensure that every syllable is attached to at least one tone, and every obligatory tone attached to at least one syllable.

We start with a relatively simple theory and consider a number of phenomena which threaten to complicate it. Every chapter, except for the

second, focusses on a generalization about the structure of English tonal morphemes which is attractive but questionable. I will either attack the hypothesis or defend it against unnecessarily complex alternatives.

In Chapter 1, I defend the hypothesis that intonational tunes are tonal morphemes. I begin by describing the phonological representations and types of tone association rule which are assumed in autosegmental phonology. Since some of these rules refer to the prominence feature specifications of syllables, prominence features are briefly discussed. A few tone association rules are postulated for English, and their application to three English tonal morphemes is demonstrated. Finally, I discuss the implications and appropriateness of the term *tonal morpheme*.

Since the distribution of the tones of a morpheme across the syllables of a phrase depends on the prominence of those syllables, it is crucial that our descriptions of phrasal prominence be correct. The specifications of prominence features are assigned by accentuation rules, which apply before the tone association rules. In the second chapter I support the analysis of prominence assumed in Chapter 1 by showing that it offers a natural account of the phenomenon of variable stress, or stress shift, as in the word *afternoon*. I examine three alternative accounts of variable stress, and show that they make incorrect predictions, or none, where mine makes correct ones.

Chapter 3 is devoted to obligatory prenuclear tones. In all the tunes discussed so far, the nucleus tone is the first obligatory tone in the tonal morpheme. Are there any English tunes with an obligatory tone before the nucleus tone? The Low Bounce and the Long Jump of O'Connor and Arnold have been identified as tunes of this kind. I argue that these tunes do not contain obligatory prenuclear tones, but that there is at least one other tune, O'Connor and Arnold's Switchback, which does.

The subject of the fourth chapter is the fifth tone. The hypothesis under attack in this chapter is that four tones suffice to model the tunes of English. All the well-known levels approaches to intonation use four levels, and many contour approaches can be expressed in terms of three levels, but my theory uses five tones. In this chapter we justify the five tones and consider how other theories get by with fewer tones.

In the fifth chapter, I argue for the abandonment of the hypothesis that

there are no accent tones following the nucleus tone of a tonal morpheme. The tune which necessitates this move is one traditionally described as a calling contour. The representation which I propose for this tonal morpheme also requires us to abandon the assumption that the nuclear syllable is always the last accented syllable in a phonological phrase. This in turn suggests that accentuation rules need information about the tonal morpheme attached to the phonological phrase on which they are operating. This hypothesis is supported by evidence from the intonation of stock auctioneers.

Chapter 6 is a contribution to the levels versus contours debate. In it I criticize the hypothesis that what I call tonal morphemes contain-- or are-- contour tones, and defend the hypothesis that tonal morphemes consist entirely of level tones, or pitch levels, against the attacks of contourists.

In the seventh chapter and the next I take up two complementary kinds of apparent exception to the mononuclear tune hypothesis, which is that each tonal morpheme contains one and only one nucleus tone. Chapter 7 deals with claims that some tonal morphemes must contain more than one nucleus tone. This complication is shown to be unnecessary and incapable of explaining certain facts about so-called compound tone units which are explained by a simpler theory consistent with the mononuclear tune hypothesis.

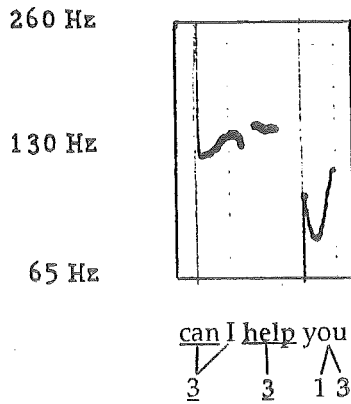
Chapter 8, the last, is devoted to the opposite type of exception-- a tonal morpheme with no nucleus tone. Liberman and Sag propose their Contradiction Contour as such a tune. Their treatment and the treatment by other linguists of this tune is discussed in order to determine the correct representation in terms of our theory. I conclude that this tune does indeed lack a nucleus tone, and that it is exceptional in other ways, as well. In particular, its accent tone differs from those in other tonal morphemes. The tune is also shown to occur with utterances which are not contradicting.

Finally, this seems to be an appropriate place for a few words on the phonetic data used in the following pages. This thesis contains a number of visual records of the pitch variation in English utterances. These were produced by a chart recorder connected to a laryngograph. The laryngograph, whose history and operation are described by Fourcin and Abberton (1971), is a

machine which measures impedance variation in a low voltage current passing between two small metal plates. If these plates are placed on the neck on either side of the larynx, the impedance will vary with the degree of opening of the vocal folds. When the vocal folds vibrate, or open and close rapidly, the machine will produce a signal which varies with the frequency of vibration. This is derived by measuring the period of vibration and calculating the negative of its logarithm. When this signal is displayed, sounds whose fundamental frequency is an octave apart are equidistant on the y-axis. This axis is therefore equivalent to a logarithmic fundamental frequency scale.

A logarithmic fundamental frequency scale is often described as a pitch scale, pitch being a perceptual, or auditory, term, rather than a physical one. As an indication of the way humans hear differences in pitch, it is certainly more accurate than a linear frequency scale. Because of this, it is preferable to the linear frequency scales used in other works on intonation, such as Liberman 1978, Bing 1980, and Pierrehumbert 1980, but it can be called a pitch scale only if we restrict the term to musical pitch. As Stevens and Volkman (1940:334) point out, "octaves in different parts of the musical scale do *not* sound like equal intervals of pitch". A true pitch scale would be calibrated in units such as mels, a mel being defined as 1/1000 of the pitch of a 1000 Hz (cycles per second) tone. Over a range of only two octaves, however, the difference between a mel scale and a logarithmic fundamental frequency scale is minimal.

The machines have been adjusted so that an octave is represented by 12 mm, and a semitone by 1 mm on the chart recorder paper. These records have been enlarged 41% by photocopying, so that an octave is represented by about 17 mm in the diagrams in this work. The horizontal scale is time, of course. Since four seconds of speech takes up about 68 mm on the chart recorder paper, one second of speech occupies approximately $(68 \div 4 \times 1.4 =)$ 24 mm in the enlarged traces. The lower and upper boundary lines in the following diagram represent 65 and 260 Hz respectively, and 130 Hz is halfway between them. The top and bottom lines are therefore two octaves apart. The fainter horizontal lines are approximately two semitones apart.



Because the laryngograph displays are reproduced unretouched, they may sometimes be difficult to interpret, for a number of reasons. It is necessary to use words with voiceless consonants so that syllable boundaries can be seen. This has an unfortunate consequence. It produces a spiky display which reflects sudden changes-- irrelevant for our purposes-- in the state of the vocal cords as they stop and start vibrating. Larger perturbations are caused by interference and by the tracer returning to a baseline between voiced segments. The difference between the pitch lines and those which represent "noise" is less obvious on the photocopied displays than on the originals, but the fact that the noisy lines are fine and vertical, or nearly so, makes it relatively easy to distinguish them from the lines which correspond to pitch changes. I trust the reader will agree that in spite of its disadvantages, a "dirty" display is preferable to one that has been "cleaned up".

The utterances were produced by the author, a Scottish-born speaker of Canadian English. It seems to be true that, as Cruttenden (1986:174) says, "the majority of English dialects share a very similar intonational system", in particular the standard dialects of Australia, New Zealand, Canada, the United States, and southern England. Insofar as these dialects share the characteristics which I discuss, my conclusions apply not just to Canadian English but to all of them.

Chapter 1

Tonal Morphemes

This work is concerned with a number of basic questions about English intonational units. For clarity, if for no other reason, these questions must be framed in terms of a particular theory of intonation. Thus certain assumptions about the form of intonational units are taken for granted in order that other questions may be investigated. The Introduction contains a brief outline of the theory presupposed in the following chapters, but since it is not quite the same as any published theory, I will discuss certain aspects in more detail in this chapter. Some aspects will be justified here, and others in subsequent chapters. I will not attempt to offer a comprehensive defence and comparison with alternative theories before proceeding, for three reasons. First, it would occupy so much time that none would remain to deal with the questions which I wish to discuss. Second, if we were to postpone investigation of these questions until the correct representation of English tunes was established, we would have to postpone it indefinitely, because certainty in such matters is unattainable. Finally, investigating such questions typically turns up evidence which bears crucially on the choice between competing theories.

Let us suppose that English has an inventory of tunes, and that every well-formed phonological phrase is associated with one tune, or tonal morpheme. I call these units tonal morphemes because I suppose that they are not phonological units and that their immediate constituents are tones. I will discuss the implications of this term at greater length in section 4. I take tones to be phonological units, and I argue in Chapter 4 that the number of tones required to characterize English tonal morphemes is five. I use integers to represent the tones, calling the lowest one tone 1 and the highest tone 5. The higher the number, the higher the pitch of the tone, *ceteris paribus*. The

identity of the tone (or tones) linked to a syllable is an important determinant of that syllable's pitch, but it is not the only one. Some of the others will be discussed in section 4.

A phonological phrase is a sequence of syllables. A tonal morpheme is a sequence of tones. The two sequences need to be attached to one another, so that every syllable is linked to at least one tone, and is thus pronounceable. The linking is done by tone association rules. I expound these in section 1, outlining the conventions of autosegmental phonology which they presuppose. Because the tone association rules refer to prominence features, I will introduce these features in section 2, before demonstrating the application of the tone association rules to a selection of tonal morphemes in section 3.

1. Autosegmental Phonology

1.1. Background

Autosegmental phonology arose out of dissatisfaction with certain aspects of the phonological theory presented in Chomsky and Halle 1968. Chomsky and Halle assume that a phonological representation consists of a linear sequence of units, each unit being a complex symbol consisting of a set of specifications for a set of binary features.

k	æ	t
-syll	+syll	-syll
-son	+son	-son
+high	-high	-high
-ant	-ant	+ant
-cor	-cor	+cor
-voice	+voice	-voice
-cont	+cont	-cont
<i>etc.</i>	<i>etc.</i>	<i>etc.</i>

These complex symbols are either segments or boundaries. Phonetic symbols are abbreviations for particular combinations of feature specifications. A phonological representation is a single sequence of segments (and boundaries).

In the early 1970s people began suggesting that this concept of what a phonological representation is like was inadequate. The facts that led linguists to this conclusion, and to the modification of phonological theory, had to do mostly with pitch in African languages. In many African languages pitch-- the occurrence of different tones-- can distinguish one word or morpheme from another. That is, they are tone languages. The relevant phenomena include the following:

- a) contour tones in level-tone languages,
- b) stability of pitch features-- i.e. the persistence of a tone despite the removal of the vowel which appeared to be carrying it,
- c) floating tones-- morphemes consisting only of tones,
- d) purely tonal regularities of morpheme structure.

The solution proposed by Goldsmith (1976) and others to these problems was a phonological representation consisting not of one but of two linear sequences of units-- separate or autonomous sequences of segments. This is the autosegmental solution. The different subrepresentations are called tiers. In Goldsmith's terminology both of these tiers are autosegmental: autonomous and segmental. To say that they are independent is not to say that they are not connected. The segments on one autonomous tier are linked to those on a second tier by means of association lines. I will assume for the sake of simplicity that tonal segments are associated with syllables on the non-tonal tier. There is nothing to prevent one syllable being associated with more than one tone. Conversely, nothing prevents a tone being associated with more than one syllable.

Goldsmith does not say that tone is always autosegmental. This is one way in which this term differs from suprasegmental, which is traditionally used of certain features, e.g. loudness, length, pitch-- that is, phonetic criteria are used to decide what is suprasegmental. The criteria for deciding whether a feature is autosegmental are phonological: how does it behave in the phonological rules of a particular language? So a feature like tone can be autosegmental in some languages and not autosegmental in others.

Goldsmith does not say that tone is the only feature that can be

autosegmentalized. Indeed he suggests that nasality is on a separate tier in South American languages like Terena and Guaraní. And vowel harmony-- manifested in a feature such as Tense, or Raised, or Advanced Tongue Root-- is also amenable to an autosegmental approach.

Subsequent developments in autosegmental theory have led linguists to split up the non-tonal or basic tier, so the number of tiers has increased, but not in the manner prefigured in Goldsmith's work. The theory has been extended in a direction which is not at all obvious from a reading of early autosegmental phonology. McCarthy's (1981) work on non-concatenative morphology in Semitic languages was very influential in bringing about these changes. The best-known extension is three dimensional phonology.

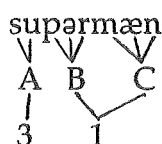
Goldsmith proposes that a particular feature will be autosegmental in a particular language-- i.e. represented on a separate tier from the usual sequence of segments-- if it undergoes certain kinds of phonological rule in that language. So some languages have one tier, some have two tiers, some have more than two. In three dimensional phonology as described by Halle and Vergnaud (1980), all languages have a number of tiers, one of which consists of syllables.

1.2. Representations

I take it for granted that a phonological representation consists of several parallel sequences, or tiers, of linguistic units. This is the assumption which distinguishes autosegmental or three-dimensional phonology from other kinds of phonology. Three of the sequences are especially relevant here:

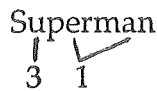
- (i) a sequence of phonemes;
- (ii) a sequence of syllables;
- (iii) a sequence of tones.

I will assume that the syllable tier is central, and that both phonemes and tones are attached to syllables, as in the following diagram.



The relationship between the phonemes and syllables is oversimplified in this

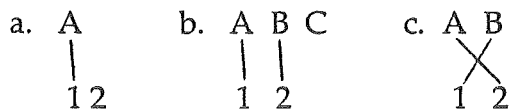
diagram, in that it omits the hierarchical structure which syllables appear to have. In this work the diagrams will be further simplified by omitting the syllable tier entirely, and replacing the phoneme tier with a representation in standard orthography. So a diagram like the following is to be understood as an abbreviation of the one above.



Our phonological representations must be constrained in various ways, because we want our theory to characterize some representations as impossible in a natural language. Goldsmith suggests we do this by requiring that phonological representations conform to a Well-Formedness Condition, and proposes this three-part condition.

- (a) Every tone is associated with some syllable.
- (b) Every syllable is associated with some tone.
- (c) Association lines must not cross.

The first of these rules out representation a below, the second rules out b, and the third rules out c.



1.3. Rules

The rules responsible for linking tones and syllables with one another go by various names. I will call them tone association rules. We can distinguish three kinds. The first kind are initial tone association (ITA) rules. They come into operation first, before any syllables and tones have been linked. They are language-specific, but they all appear to conform to a small number of schemata. Then there are well formedness (WF) rules. These are supposed to be language-universal, and come into operation after the other types of rule. Rules of the third kind are able to add, change, and delete association lines which are already present in a representation. Goldsmith's "flop rule" is an example of this kind of later tone (LT) rule. The output of the ITA rules and the WF rules is the input to the LT rules.

1.3.1. Initial Tone Association rules Clements and Ford (1979:181) distinguish three types of ITA rule, as follows.

1. Associate a designated tone of the tonal string with the tone-bearing unit that bears the accent.
2. Associate a designated tone of the tonal string with the leftmost (or rightmost, penultimate, etc.) tone-bearing unit.
3. Associate tone n of the tonal string with tone-bearing unit m (counting from the left). (Usually $n = m = 1$.)

In this chapter I will demonstrate the application of the first two types in English. To my knowledge no one has proposed that the third type is also used, but I would not rule it out.

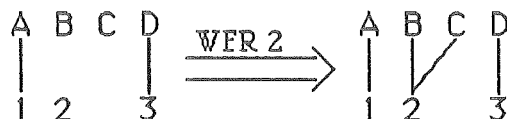
1.3.2. Well-Formedness rules These rules bring ill-formed representations into line with the Well-Formedness Condition. They come into operation after the operation of the ITA rules and make any additions or deletions necessary to make the representation well-formed. Rules are required because there may be more than one way to rectify particular ill-formed representations. A constraint such as Goldsmith's WFC provides no guidance on preferred alternatives. Clements and Ford (1979:183-186) propose three Well-Formedness rules (or conventions, as they call them). They refer to facts which suggest that the first part of Goldsmith's WFC is not right. In some languages, some tones may remain unattached. Since Clements and Ford's rules are meant to be universal, they leave tones stranded in some positions, to be taken care of by language-specific rules. I give here their three rules, which I shall assume to be sound. Because Clements and Ford want their rules to work for languages in which features other than tonal ones are autosegmental, they use the terms *melodic element* and *melody-bearing element*. In the following statement I have replaced these with the less general terms *tone* and *syllable*, respectively. Letters represent syllables, and numerals represent tones.

Well-Formedness Rules

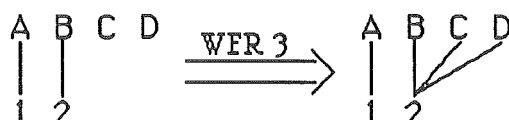
1. If there are several unassociated tones and several unassociated syllables, the former are associated with the latter from left to right.



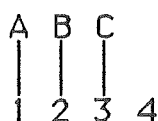
2. If after the application of the first convention, there remain one unassociated tone and one or more syllables, the former is associated with all of the latter.



3. If all tones are associated and if there are one or more unassociated syllables, all of the latter are assigned the tone associated with the syllable on their immediate left if possible. That is, tone spreads to the right.



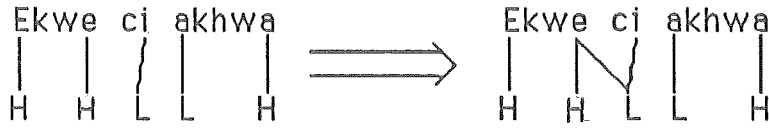
As I mentioned above, these rules make no provision for the automatic association of an unassociated tone with an associated syllable. Only by a language-specific rule can the floating tone 4 in the following representation be anchored to a syllable.



It will become clear that such a rule is required for English. One of our goals will be to determine the precise form of this rule-- or rules. For the moment we will assume that English has a fourth Well-Formedness Rule which associates tone 4 in the above example with syllable C. This leaves several questions unanswered, one of which will be taken up in section 3.3 below.

1.3.3. Later Tone rules Rules of the third type are more like ordinary phonological rules, such as rules of assimilation. According to Goldsmith (1976:25), Igbo has a tonal association rule which operates as follows. If the final syllable of a noun phrase subject (e.g. *Ekwe*) is normally high, it becomes falling

if the following syllable (e.g. *ci*) is low. This can be explained as the result of a rule which associates the last syllable of the subject with the tone attached to the following syllable. The final syllable of *Ekwe* is thus associated with both a high tone and a low tone, in that order. The pitch of the syllable is high at the beginning and low at the end-- falling.



I follow Goldsmith and others in assuming that LT rules, as the name implies, apply after the WF Rules.

2. Prominence

Tone is the phonological correlate of pitch (an auditory, or perceptual, term) and fundamental frequency (an acoustic, or physical, term). Prominence is the phonological correlate of loudness (an auditory term) and intensity (an acoustic term). Prominence is therefore a cover term for phonological features which make some syllables stand out more than others. The terms *stress* and *accent* have traditionally been used for these features, but they have not been used consistently. In fact, they have probably been used less consistently than any other pair of linguistic terms one can think of.

2.1. Three kinds

I assume that prominence features are features of syllables, and that the syllable is a fundamental linguistic unit, and therefore indefinable. There does not seem to be any language in which all syllables are equally prominent. In some languages there is apparently only a single binary feature, that is, a two-way distinction. In those languages, each syllable of an utterance is either prominent or not, stressed or not-- whatever term one wishes to use. This may do for some languages, but for English phrases we need to recognize three kinds of prominence. To this end we use three binary features: [\pm strength], [\pm accent], and [\pm nucleus]. Because the features are hierarchically related rather than cross-classifying, these three features yield only four levels of prominence. All four are illustrated in citation utterances of the following phrases.

	Afternoon tea					A confidential letter came							
Nucleus	-	-	-	+	Nucleus	-	-	-	-	-	+	-	-
Accent	+	-	-	+	Accent	-	+	-	-	-	+	-	-
Strength	+	-	+	+	Strength	-	+	-	+	-	+	-	+

The most prominent syllable is the one with the greatest number of plusses beneath it. Let us look briefly at each of these two-way distinctions.

2.1.1. Strength The first distinction is between **strong** [+strength] and **weak** [-strength] syllables. All syllables in a phonological phrase are either strong or weak. Strong syllables stand out more than weak ones. The alternation between strong and weak syllables is responsible for the rhythmic nature of spoken English. These rhythms are exploited in many kinds of poetry, where the pattern of strong and weak syllables has to fit-- more or less-- into a framework. Different verse forms require different arrangements. The first line of a limerick, for example, has a specified pattern of strong and weak syllables, indicated by dashes and dots respectively in the following.

. - . . - . . -
There was a young fellow called Hyde.

For a second example we can use the same sentence. If this sentence were not the first line of a limerick, it would be spoken differently. In particular, the strong and weak syllables would be differently arranged, possibly as follows.

. . . - - . - -
There was a young fellow called Hyde.

Dots and dashes are appropriate for introducing this type of prominence, but they are not very convenient for transcriptions. For this reason, strong syllables in the transcriptions in this work will be marked with the IPA stress diacritic: a small vertical stroke above the print line at the beginning of the syllable. Weak syllables will be left unmarked. The two examples given above are represented as follows.

1. There 'was a young 'fellow called 'Hyde.
2. There was a 'young 'fellow 'called 'Hyde.

Utterances of the following phrases out of the blue would most likely have the indicated pattern of strong and weak syllables. The first has four strong syllables: *con, den, let, came*. The second has three: *af, noon, tea*.

3a. A 'confi'dential 'letter 'came.

4a. 'After'noon 'tea.

A couple of remarks on the phonetics of strength would not be out of place here. Strength is related to intensity, vowel quality, and duration. I will ignore the first two, because they cannot be observed by means of a laryngograph. Strength is one of the factors which determines the length of a syllable. The length of a strong syllable varies inversely, *ceteris paribus*, with the number of weak syllables following it in the same foot, the latter being a phonological unit containing one strong syllable. The formula $a / x + 1$ gives a very good approximation, where a is the length of a syllable in isolation, and x is the number of syllables following it. The number of weak syllables before a strong syllable has no effect on its length. The length of a weak syllable does not vary in this way. Its length is about the same as a phonologically comparable strong syllable followed by three weak syllables. To present the evidence for these claims would take us beyond the scope of this study, but I mention them here to give the reader a better idea of what I mean by strength.

2.1.2. Accent The second distinction we make is between **accented** [+accent] syllables and **unaccented** [-accent] ones. This distinction is significant only for strong syllables. All weak [-strength] syllables are unaccented. Some strong syllables are accented, and some are not. Accented syllables are more prominent than unaccented syllables. Consider again a citation utterance of this sentence:

3a. A 'confi'dential 'letter 'came.

I have said that both *con* and *den* are strong, but to native speakers of English these two syllables are not equally prominent. In this utterance *con* sounds louder than *den*. This is because *con* is accented and *den* is not. The syllable *den* is also less prominent than the following strong syllable, *let*, because *let* is accented. We indicate accented syllables by underlining them:

3b. A 'confi'dential 'letter 'came.

So this phonological phrase has four strong syllables, and two of these are also accented. Our other example, afternoon tea, also has two accented syllables.

4b. 'After'noon 'tea.

The first and last syllables stand out more than the third. The first and last syllables are not, however, equally prominent.

2.1.3. Nucleus This brings us to the third distinction, which is between nuclear [+nucleus] and non-nuclear [-nucleus] syllables. All unaccented [-accent] syllables are non-nuclear. Nuclear syllables are more prominent than non-nuclear ones. Each phonological phrase contains one and only one nuclear syllable. In written transcriptions we mark nuclear syllables by double underlining, but this is not convenient in typed or printed work. In this work, nuclear syllables will be in boldface as well as underlined.

3c. A 'confi'dential 'letter 'came.

4c. 'After'noon 'tea.

The transcription system is redundant. An accented syllable is necessarily strong, but it is marked strong anyway. A nuclear syllable is necessarily accented, but it is marked accented as well. In the following pages, I will sometimes omit the strength marks and the underlining when the location of [-nucleus] syllables is irrelevant to the discussion.

2.2. Prominence and Phrasing

Both prominence and intonation in English are closely related to phrasing. One of the differences between listening to a language one does not know and listening to a language one does know is that in the latter one hears gaps which do not exist physically. Clearly what enables one to do this is one's knowledge of the language. If we wish to characterize that knowledge, our theory of language must account for those gaps. Postulating entities such as morphemes and words and phrases and sentences allows the linguist to account for some of these gaps-- they correspond to the boundaries of these grammatical units. There appear to be gaps of another kind which correspond in large part to the commas and full stops of the language in its written form. These gaps sometimes coincide with physical gaps, or silences, in speech, but often they do not. Neither is there a consistent relationship between these gaps and syntactic structure. The same sentence can be spoken with gaps in different places, and the material between the gaps does not always correspond to syntactic units.

We can account for these gaps by postulating a phonological unit which I will call the **phonological phrase**, and supposing that native speakers perceive a pause at the end of each phrase. Where necessary, I will indicate the end of a phonological phrase with a pair of vertical lines, thus: | |.

The same unit is relevant to the distribution of prominence feature specifications. That is, phrasing is related to prominence. I will suppose the following hypotheses about the relationship to be true, at least initially.

- i) Each phonological phrase contains one and only one nuclear syllable.
- ii) There is no more than one accented syllable before the nuclear syllable of a phonological phrase.
- iii) The nuclear syllable is the last accented syllable in a phrase.

These generalizations will be explained in Chapter 2 as the consequences of the operation of a set of accentuation rules. The phonological phrase is the domain of the accentuation rules, which make syllables strong, accented, and nuclear.

I have said that phrasing is also related to intonation. In particular, I suppose that every well-formed phonological phrase is linked to one and only one tonal morpheme. The domain of the tonal morpheme-- the span of a sentence which is associated with a tonal morpheme-- is thus hypothesized to be the same as the domain of the accentuation rules. Other names for the phonological phrase, such as *intonation unit*, *tone group*, and *tone unit*, reflect their users' adoption-- not always made explicit-- of comparable hypotheses. It may turn out that the phonological phrase and what we might call the intonational phrase are in fact distinct, but the more economical assumption that they are identical has produced good results.

3. English Tonal Morphemes

I list here the tonal morphemes which I will refer to in this study, in order of appearance. The first column contains the number of the chapter in which each tune is first discussed. The second contains a representation of each tonal morpheme as a sequence of tones represented by numbers, some with diacritic features which will be explained shortly. The third column contains a name for each tonal morpheme. These names are traditional rather than consistently

form-based or function-based. Where the tonal morpheme can be identified (more or less) with a "tone group" described by O'Connor and Arnold (1973), I use the names which they use. This applies to the first eight in the list, with the exception of "Dipper", which I have invented in the spirit of O'Connor and Arnold, and "High Rise", which replaces their "High Bounce". In this chapter I will use the first three morphemes in the list to illustrate the operation of the tone association rules.

Ch.	Form	Name
1	(1)(3) <u>3</u> 1	Drop
1	(1)(3) <u>1</u> 3	Low Bounce
1	(1)(3) <u>3</u> 1 <u>3</u>	Dipper
3	(1)(1) <u>1</u> 3	Take Off
3	(3)(1) <u>3</u> 1	Long Jump
3	(1)(3) <u>1</u> 3 <u>1</u> 3	Switchback
4	(1)(3) <u>2</u>	Terrace
4	(1)(3) <u>4</u> 5	High Rise
4	(1)(3) <u>4</u>	High Level
5	(1)(3) <u>3</u> 2	Calling Tune
5	(1)(3) <u>3</u> 2	Conductor Tune
8	[33] <u>1</u> 3	Contradiction Contour

3.1. The Drop

I will begin with the tonal morpheme which is apparently the most common in the standard accents of English. This morpheme contains two obligatory tones: 3 followed by 1. The first of these gets attached to the nuclear syllable of its phonological phrase. We mark nuclear syllables with boldface type and underlining, and we will use the same thing to mark those tones which are to be linked with nuclear syllables. *Nuclear tone* would be a fitting name for such a tone, but the term is so widely used with a different meaning, I feel obliged to avoid it. I will refer to the boldface underlined tone of a tonal morpheme as a **nucleus tone**.

Before I introduce the optional tones of this morpheme, let us see how the

tone association rules work with this tone when the nuclear syllable is the first one in the phrase, taking a phonological phrase consisting of the word *Superman*. Our intonation rules require as input both a particular tonal morpheme and a phrase marked for prominence. The accentuation rules have already operated, and for each syllable of the phrase we know whether it is strong or not, accented or not, and nuclear or not. In our prominence transcription of *Superman*, the first and third syllables are strong, and the first syllable is both accented and nuclear.

5a. 'Super'man

3 1

First of all, the language-specific tone association rules, the ITA rules, go to work. The rule which attaches nucleus tones is an example of the first type of ITA rule distinguished by Clements and Ford (1979), which I repeat here:

ITA Rule Type 1

Associate a designated tone of the tonal string with the tone-bearing unit that bears the accent.

In English, of course, there may be more than one accented syllable, or tone-bearing unit, in the domain of the rule. I have proposed that English has two kinds of accented syllable-- nuclear and nonnuclear-- and that a phonological phrase may contain one of each. Accordingly, there will be two English ITA rules of the first type. The first may be stated as follows:

English ITA Rule 1a

Associate a nucleus tone with the nuclear syllable.

The output of the this ITA rule looks like this:

5b. 'Super'man

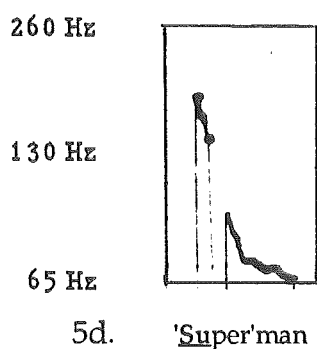
3 1

The Well-Formedness rules, discussed in section 1.3.2 above, now come into operation. The second of Clements and Ford's universal WF rules will link the unassociated tone with the two unassociated syllables to produce a well-formed phrase in which no syllables or tones are left without partners, as follows.

5c. 'Super'man

3 1

This representation makes a rough prediction about the pitch contour of such a phrase. We can work this out using a kind of musical stave with a line for each of our tones. I say a rough prediction because the height of a tone is only one of the factors which determine the actual frequency and pitch of a syllable or part thereof. Some of the other factors will be discussed in section 4.1 below. At this point it is enough to note that, because of the phenomenon of downdrift, the lines of our stave will not be horizontal. They will have to slope downwards. We therefore predict that the second syllable will be much lower than the first, and slightly higher than the third. The predicted contour can be compared with a laryngograph display of an actual utterance of the phrase.



The laryngograph is designed to produce a signal only when the vocal folds are vibrating. When the folds are parted for breathing or the production of voiceless sounds, there is no trace. Since the first sound of *Superman* is voiceless, we see nothing corresponding to the /s/. Then there is a line for as long as the /u/ lasts, then a gap during the closure for the voiceless bilabial stop, and then a line corresponding to the remaining sounds, all of which are voiced. In our utterance of *Superman* with the 31 tonal morpheme, we see that the last two syllables are considerably lower than the first. We can see that it takes some time for the frequency to drop after the /p/, so that in phonetic terms the pitch of the second syllable does not reach tone 1 until mid-syllable. Such delays are typical of speech, as opposed to singing. As the word *contour* suggests, pitch contours do not normally have right angles.

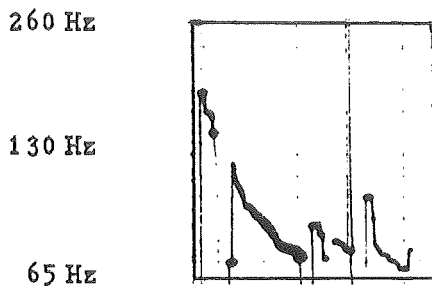
The initial tone is identified as 3 rather than 2 or 4 because we need one tone between it and the lowest tone, and no more. The tonal morpheme which

requires the 2-tone will be discussed in Chapter 4.

Let us now consider a longer phonological phrase, *Superman couldn't do it*, with the first syllable nuclear.

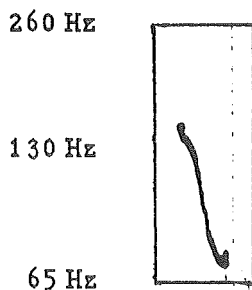
- 6a. 'Super'man 'couldn't 'do it
 3 1

If we use the same tonal morpheme for this phrase, the 3-tone gets linked to the nuclear syllable by English ITA Rule 1, and the 1-tone gets attached by WF Rule 2 to all the following syllables. This produces the desired result.



- 6b. 'Super'man 'couldn't 'do it

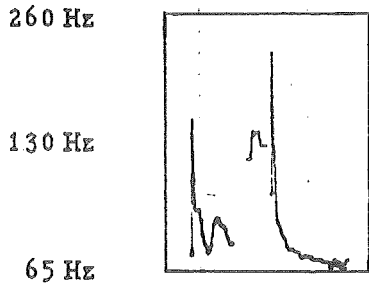
If the phrase consists of only one syllable, that syllable will be high at the beginning and low at the end. It is attached to both the 3-tone and the 1-tone, from left to right.



7. 'Sue
 3 1

None of the three universal WF rules postulated by Clements and Ford will in fact attach the 1-tone to a syllable which is already associated with a tone. This will have to be done by a fourth WF rule which is not universal. We do not have to specify the details of the rule to deal with this instance, because there is only one syllable to which the floating tone can be attached. I shall return to the details of this rule in section 3.3 below.

In the three examples used so far, the nuclear syllable has also been the first syllable. What happens if it is not? Consider the phrase *He's a Superman* with the first two syllables weak.



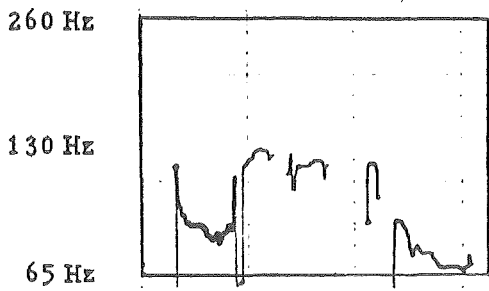
8a. he's a 'Super'man

The first two syllables are relatively low in pitch. Let us say they are attached, like the last two syllables, to a 1-tone. Phonetically, of course, the last two syllables are definitely lower, but this can be attributed to downdrift. So this phrase will be represented as follows.

8b. he's a 'Super'man
 1 3 1

This means we need to amend our representation of this tonal morpheme by adding a 1-tone at the beginning: 131. There is a difference, though, between the initial 1-tone and the final one. The final one must get attached to some syllable, even if the only syllable available is already associated with another tone. This is what happens with single-syllable phrases like *Sue*. The initial 1-tone, on the other hand, is optional. It is attached only to free syllables. I will indicate this by putting parentheses around this tone in our representation: (1)31.

Thus far all our examples have contained only one accented syllable-- the nuclear syllable. Let us now consider what happens if there is an accented syllable preceding the nuclear syllable of the phonological phrase, as in *he's 'built like 'Super'man*.

9a. he's 'built like 'Super'man

The first syllable is relatively low, attached to the initial 1-tone. The second and third syllables are higher, both at about the same pitch as the nuclear syllable. Since tone spreads to the right rather than the left, we cannot have the second and third syllables attached to the nucleus tone. What we require is another 3-tone, after the optional 1-tone, which gets attached to a non-nuclear accented syllable, if there is one. In other words, it is optional in the same way as the initial tone, and we can indicate this with parentheses. I will call this sort of tone an accent tone. Accent tones are linked to nonnuclear accented syllables by English ITA Rule 1b.

English ITA Rule 1b

Associate an accent tone with the nonnuclear accented syllable.

Accent tones will be marked in the same way as nonnuclear accented syllables--with single underlining. Our complete representation of the first tonal morpheme is therefore (1)(3)31, and the phrase illustrated above is represented in the following way.

9b. He's 'built like 'Super'man
 1 3 3 1

The tone association rules which associate accent and nucleus tones with accented and nuclear syllables exemplify the first of the three types of ITA rule distinguished by Clements and Ford, which for our purposes may be stated as follows.

ITA Rule Type 1 (Revised)

Associate a marked tone with a syllable marked in the same way.

Here we find justification for treating prominence in terms of a system of binary features rather than as a multivalued feature in the manner of Chomsky and Halle (1968) or as a relational feature of constituent structure in the

manner of Liberman and Prince (1977). When we postulate a feature we predict that there will be at least one kind of rule which applies only to units which have that feature. In this case we can say that all and only the syllables which are [+accent] get associated with tones by means of ITA Rule Type 1. A theory in which prominence is multivalued can make such a generalization only in a clumsy way, and the same seems to be true of a theory in which prominence is relative, defined in terms of a hierarchical structure.

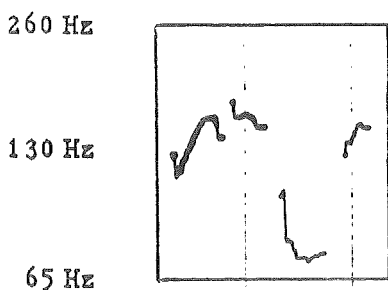
3.2. The Low Bounce

All English tonal morphemes--even the odd ones-- contain at least one tone which gets associated by an ITA rule of the first kind. Only some English tonal morphemes, however, include a tone which gets associated by the second kind of ITA rule, repeated here for reference.

ITA Rule Type 2

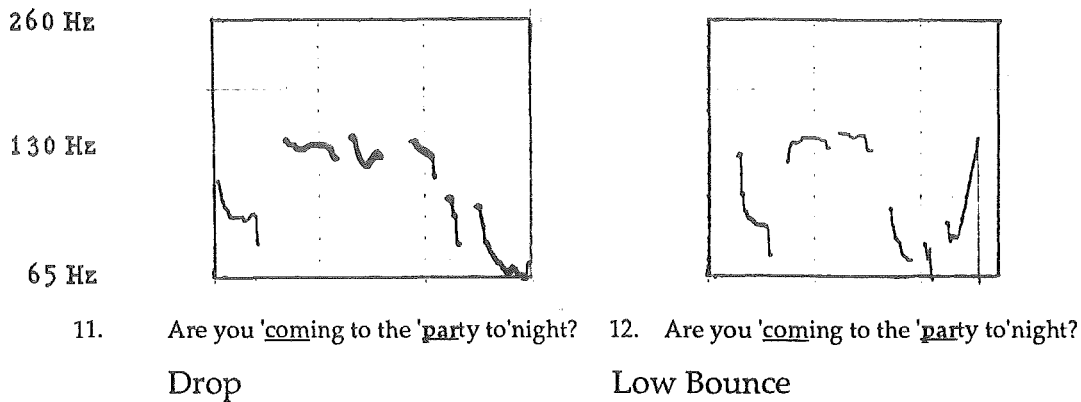
Associate a designated tone of the tonal string with the leftmost (rightmost, penultimate) tone-bearing unit.

According to the present theory, most such tones in English become attached to the last, or rightmost, syllable in a phonological phrase. I will refer to them as **end tones**. One is contained in the tonal morpheme which O'Connor and Arnold (1973) call the Low Bounce. In order to help the reader recognize this tune, I will say something about its use. It is frequently used in nonfinal phrases, such as the first clause of *As soon as you see him, tell him I called*. In Standard English English, it is apparently the normal way of asking yes-no questions like *Seen the Times leader today?* It is also commonly used in speaking to children: *Use your hankie*.



10. 'use your 'hankie

The optional, or prenuclear, tones of this tune are exactly the same as those of the first tune discussed. This is clear if we compare the same phrase spoken with each tune.



The Low Bounce also has the same obligatory tones as the Drop, but their order is reversed: 13 instead of 31, and the 1-tone is nuclear. This morpheme differs in another important way from the Drop. In the Drop, the second obligatory tone is attached to the syllable immediately following the nuclear syllable and to any syllables between it and the end of the phrase. Thus the last three syllables of example 11 above are low. With the Low Bounce, however, the second obligatory tone, in this case a 3-tone, is attached only to the final syllable of the phrase. All the syllables between the nuclear syllable and the final syllable are low, that is, attached to the same 1-tone as the nuclear syllable. This can be seen in 12 above.

Our Well-Formedness rules will produce the correct result if the second obligatory tone is attached to the final syllable before they come into effect. This will be done by English ITA Rule 2, which is an example of the second type of ITA rule.

English ITA Rule 2

Associate an end tone with the rightmost syllable of the phrase.

In written transcriptions end tones can be marked with two short vertical strokes under the numeral, since we indicate the end of a phonological phrase with two vertical strokes. In this study, however, we will represent such tones by means of an outlined numeral, thus maintaining a visual resemblance between such tones and the mark for the end of a phrase. Our representation of the Low Bounce will therefore be (1)(3)13, and example 12 represented as

follows:

12a. Are you 'coming to the 'party to'night?

We introduced an ITA rule of the second kind in order to account for the difference between the mapping of a *Drop* and the mapping of a *Low Bounce* onto postnuclear syllables. Apart from phrases in which the nuclear syllable is final, the obligatory part of a *Low Bounce* is not a mirror image of the obligatory part of a *Drop*. It is not possible to account for this difference in a framework where the difference between 31 and 13 is that one is a falling nuclear tone and the other a rising one. Crystal, for example, makes no attempt to account for this difference. The reason is that he does not recognize its existence. According to him (1969:223), the tail, which consists of the postnuclear syllables of a phrase, normally and most frequently "continues the direction of the nucleus in an unbroken fall or rise". This is what a theory such as Crystal's predicts-- that a rise will be a mirror image of a fall. We have just seen that this is not true. If we wish to explain why Crystal and others make the sort of claim just quoted, it seems to be a result of concentrating on the compressed or contracted forms of tunes, as they are manifested in a single syllable: a rise is more or less a mirror image of a fall in a monosyllabic phonological phrase. This is like trying to work out the phonological structure of *cannot* from an examination of *can't*. The shorter the sequence over which a melody is manifested, the harder it is to work out the structure. Probably this approach is primarily a consequence of the belief that these were indivisible phonological units-- tones, rather than sequences of tones.

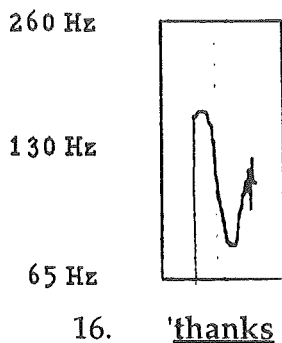
3.3. The Dipper

I turn now to the last of the three tonal morphemes to be introduced in this chapter. All of these tunes contain only two different tones: a high one represented by 3, and a low one represented by 1. All three have the same optional tones: (1)(3). The two tunes discussed so far have two obligatory tones. The third tune has three obligatory tones: 313. O'Connor and Arnold (1973) do not deal with it, but Halliday (1967) does, identifying it as a broken Tone 2 with a high head. His Tone 2 comes in two varieties, straight (neutral) and broken,

and each of these occurs with two kinds of heads-- high (neutral) and low. Virtually all of Halliday's examples are queries. I give some here, converted into my transcription system.

13. 'is your 'flat 'furnished 'by the 'university (p. 29)
14. do you 'want a 'biscuit (p.44)
15. 'are there (p. 49)

It is also used for echoes, and in New Zealand English one often hears it used with a friendly and casual *thanks*.



16. 'thanks

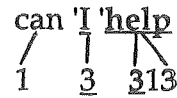
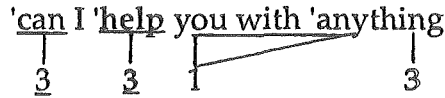
When I worked in a shop I used to use the dipper to offer assistance to customers, as in 17 and 18. Note that the first word is accented in 17 but not 18.

17. 'can I 'help you with 'anything
18. can 'I 'help

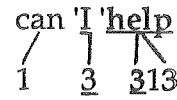
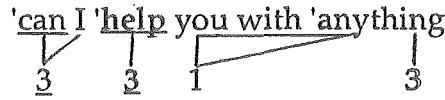
The first 3-tone gets attached to the nuclear syllable, in this case *help*. The final 3-tone, like that in the Low Bounce, gets attached to the final syllable of the phrase. We represent this tune as (1)(3)313. I give below derivations for 17 and 18.

	'can I 'help you with 'anything	can 'I 'help
	1 3 3 1 3	1 3 313
ITA #1		
	'can I 'help you with 'anything	can 'I 'help
	1 3 3 1 3	1 3 313
ITA #2		
	'can I 'help you with 'anything	can 'I 'help
	1 3 3 1 3	1 3 313

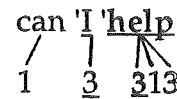
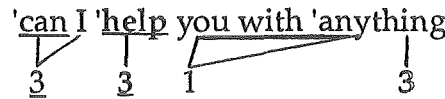
WF #2



WF #3



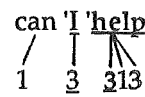
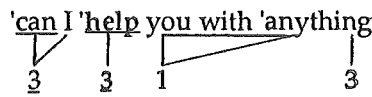
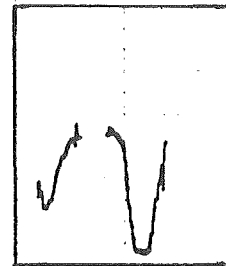
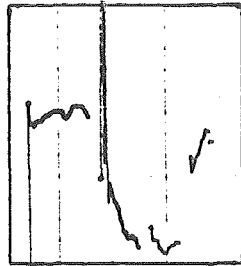
WF#4



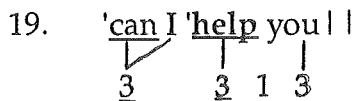
260 Hz

130 Hz

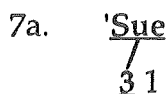
65 Hz



A more interesting phrase is 'can I 'help you. This is interesting because it provides us with information on the correct formulation of WF rule 4, which attaches tones left stranded after the first three WF rules have operated.

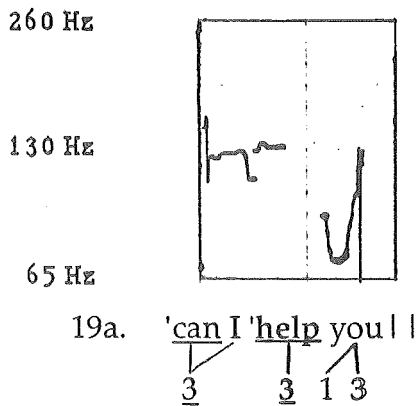


In our discussion of the Drop at the end of section 1.3.2, we saw that the 1-tone left stranded when the nuclear syllable was final had to be attached to the nuclear syllable, that is, to the syllable attached to the tone immediately left of the stranded one.



It now appears that this happens only when there is no tone to the right of the stranded tone. If we look at a trace of 19, it is clear that *help* is not low at the

end, while *you* is low at the beginning. So we attach the stranded 1-tone to the syllable associated with the tone on its right.



Let us suppose that the fourth WF rule has this form:

WF Rule 4

If after the operation of the first three WF rules there remains an unassociated tone and no unassociated syllables, associate the tone with the syllable attached to the tone on its right, if there is one. If there is no tone to the right, associate the tone with the syllable attached to the tone on its left.

This formulation may not be right. For example, it is possible that the 1-tone in this case is attached to *you* not because it is the right-hand syllable, but because it is not nuclear. That is, floating tones perhaps avoid docking with nuclear syllables, if they can. No doubt there are other possibilities, but I will not explore them here. At the end of Chapter 3 I will refer to evidence which supports the formulation above.

4. The term *tonal morpheme*

I describe intonational tunes or contours as manifestations of **tonal morphemes**. This term is not new, having been used with respect to English at least as early as 1974 by Liberman and Sag, but it is still not widely used. In the first textbook on intonation for linguists published recently (Cruttenden 1986), it is not mentioned at all. In view of this it is appropriate at this point to clarify some of the implications of the term and to attempt some justification of its

use here. I will begin with the word *tonal*, and then turn to *morpheme*. I will not discuss the question of whether they are also words, but I assume that they are lexemes, that is, they have entries in the lexicon.

4.1. Tonal

The hypothesis that the entire pitch contour of a phonological phrase represents a tonal morpheme is difficult to assess without an explication of the term *tone*. In this section I will make explicit my assumptions about the nature of tones. Relatively abstract, or phonological, characteristics are discussed in 4.1.1. In 4.1.2 I consider the more phonetic matter of the relationship between tone and pitch, paying particular attention to sources of variation in the pitch of a tone.

4.1.1. Phonology of tone First of all, I assume that tone is autosegmental. That is, given that a phonological representation consists of a number of autonomous tiers of phonological entities, tones are represented on a different tier from phonemes and syllables.

I further suppose that tones are phonological units which differ from one another in one dimension only. They are therefore not contour, or dynamic, tones, i.e. rising, falling, and so on. They are all level, or static, tones. Anderson (1978) argues that contour tones are not needed for any of the languages for which they have traditionally been used. I will argue in Chapter 6 that this applies to English, too.

I assume that there is a universal set of tones which all languages draw on. This set is part of the inbuilt human language acquisition device which a universal grammar characterizes. Many languages appear to use only two tones, but some use as many as five. In Chapter 4 I argue that five are necessary and sufficient for English, as well. Since I know of no language which requires more, I will assume that our universal grammar includes a set of five tones.

In some work on intonation in an autosegmental framework, it is assumed that the symbols for tones, like the symbols for phonemes, are abbreviations for sets of phonological feature specifications. Thus a high tone, H, is an abbreviation for the feature specifications [+high, -low]. I do not make

this assumption. One of the reasons I use numerals rather than letters to represent tones is to remind the reader of this difference. There is a simple reason for my avoidance of feature specifications. To say that tones share a feature specification is to claim that they belong to the same natural class. If two tones belong to the same natural class, we expect there to be at least one phonological rule which applies only to members of that class. One justifies a set of categories by demonstrating that this is so. I have not seen any such justification for any set of tonal features used for English. In other words, I do not use tonal features because it has not been shown that postulating such features enables us to make linguistically significant generalizations about tonal forms.

Moreover, if there is a universal set of tones, and if the number is five, this is an awkward number for a set of binary features. Two features would suffice for four tones, but three are needed for five. This is uneconomical, since three binary features could distinguish a set of eight tones. More significantly, it predicts, incorrectly, that there will be languages with six, seven, and eight tones. In conclusion, it is by no means certain that tones are the ultimate constituents of tunes, but in the absence of a justified set of tonal features, it is reasonable to proceed as if they are.

4.1.2. Tone and pitch It is a reasonable initial assumption that the tones are equidistant in terms of pitch, i.e. that they correspond to five equally spaced points on a pitch scale. The problem is that we do not have a linguistic pitch scale. The displays in this study use a logarithmic fundamental frequency scale, a musical pitch scale in which sounds an octave apart are equidistant. We shall see in later chapters that this gives fairly good results, in that the interval between tones 1 and 2 appears to be approximately the same as that between tones 4 and 5, viz. about 4 semitones at the end of a phrase. We might get better results by using a pitch scale calibrated in mels, the units of pitch used by psychologists for the study of perception. And of course, it is always possible that the best linguistic pitch scale will use neither Hertz, nor octaves, nor mels.

It is often claimed that for the most part we use only the lower half (or even the lower third) of our pitch range. "In an individual speaker", says Fry

(1979:68), "the range ordinarily employed is not more than one octave and this is located in the lower part of his total voice range". This may well be true, but it would be a mistake to conclude from this observation that a typical tone 5 syllable is no more than half way up an individual's pitch range. In the traces prepared for this work, phrase-final tone 5 syllables are about one and a half octaves higher than phrase-final tone 1 syllables. Since my pitch range is approximately two octaves, it is clear that tone 5 is normally well above the middle of the range. If most speech is relatively low in pitch, this is due largely to the infrequency of the tonal morphemes containing the highest tones, 4 and 5.

The pitch of a particular tone is not fixed, even in the speech of one person, but is determined by a variety of factors. There is no reason to expect every manifestation of a particular tone to be identical. We do not expect this of other kinds of phonological units. Not all manifestations of a particular phoneme are phonetically the same. They vary in quality, duration, loudness, and so on from one occurrence to another. Indeed, one cannot as a general rule distinguish one phoneme from all the others in a language by examination of phonetic characteristics. This is an appropriate place to discuss briefly some of the sources of variation in the pitch of a particular tone.

Some of these sources of variation are beyond the control of the speaker. Each individual has his or her own pitch range, which is physiologically determined, by the size of the larynx and other factors. The tones of a person with a high voice will be higher, tone for tone, than those of a person with a low voice. Since all the utterances reproduced in this study were produced by one person this variation will not concern us at all. A second uncontrollable source of variation will affect the displays in a perceptible manner. It seems that the pitch of a vowel is affected by the type of consonant which precedes it, and by the quality of the vowel itself. Hombert (1978) provides a discussion and references. These segmental factors do not affect the displays in a gross way, but they do account for some of the perturbation. The stopping and starting of vocal fold vibration between voiced and voiceless sounds pushes the trace off line in a more obvious way, but it is not clear whether these deviations correspond to changes in the fundamental frequency. They may reflect the

limitations of the machinery which produces the visual record. Certainly these perturbations pass unnoticed by native speakers.

Turning to controllable sources, let us begin with sources of variation which operate primarily from phrase to phrase rather than within phrases. Some differences in the pitch of a particular tone between one phrase and another are due to changes in register. That is, the speaker may use a higher or lower section of her total pitch range than normal. It is generally agreed that there is more scope for upward movement than for lower movement, because the normal conversational register is closer to the bottom of the range than to the top. Cruttenden (1986:55) claims that "both the highest and lowest levels are moved upward or downward". This is a reasonable supposition. Whether the topline and the baseline are moved by an equal amount is impossible to determine without an established scale of linguistic pitch.

In speaking the examples for the laryngograph displays in this work, I found it necessary to use a register slightly higher than my conversational register in order to get a clean display. I did this in order to avoid the interference caused by creaky voice which, as Ladefoged (1975:123) observes, "occurs at the ends of falling intonations for some speakers of English".

Another source of variation in the pitch of a tone is what some analysts call the key of a phonological phrase. Brazil et al. (1980) distinguish three keys for English-- high, mid, and low-- and claim that key is selected once for each phrase. Choice of key determines both the height of tones throughout a phrase, and the appropriateness of the phrase in a particular position in a particular discourse. According to Cruttenden (1986:54), differences of key "very often signal cohesion between intonation-groups". Brazil et al. are not explicit about the phonetics of key, but Cruttenden suggests that, unlike register, key is a matter of pitch range rather than pitch height. That is, the top line is raised for high key, but not the baseline.

Normally, neither key nor register changes are responsible for variation in the pitch of a tone within a phonological phrase. Since we will deal mostly with single phrases, these sources of variation concern us less than declination and boosting, to which I now turn.

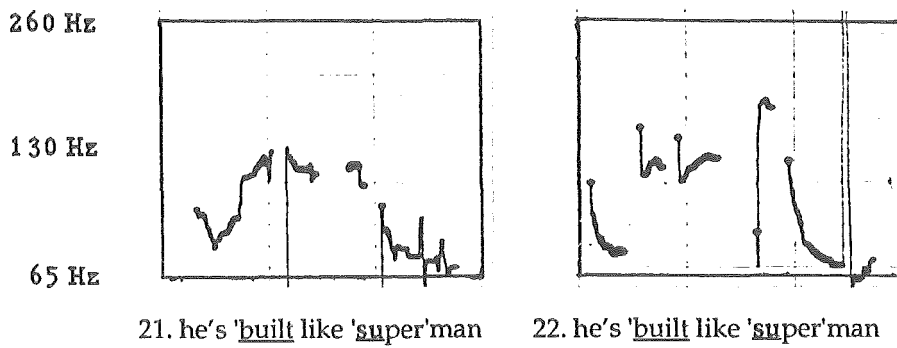
Declination, or downdrift, causes the pitch of a particular tone to be higher

at the beginning of a phrase than at the end. This phenomenon is very familiar to students of African languages, as Ladefoged (1975:229) explains:

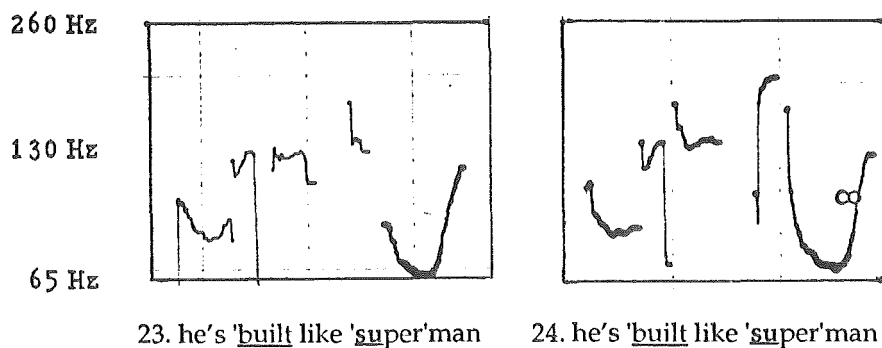
In Hausa, downdrift involves the falling of the mean pitch level throughout the sentence. Both high tones and low tones at the beginning of a sentence are higher than at the end. A high tone at the end of a Hausa sentence may even have about the same absolute pitch as a low tone had at the beginning of the same sentence.

Ladefoged talks about downdrift within sentences. I have implied that it is a characteristic of phonological phrases. It may in fact be characteristic of breath-groups. Breath force, or air flow, is normally at a maximum at the beginning of an utterance, and lowest at the end. According to Ladefoged (1975:223), pitch varies with airflow even without an increase in the activity of the laryngeal muscles. This suggests that declination is a language-independent and uncontrollable phenomenon, and its apparent presence in many other languages supports this view. On the other hand, there is evidence which suggests that declination is conventional rather than natural. As Cruttenden (1986:127) points out, it appears to be characteristic of a certain style of speech, and can be suspended. I do not intend to explore this problem here. With Pierrehumbert (1980) and others, I assume that declination affects both the baseline (tone 1) and the top line (tone 5). My identification of the initial element of the Drop tune as tone 1 reflects the assumption that the baseline declines more rapidly than the top line. We have already seen that syllables attached to the initial tone of this morpheme are both higher than final syllables and closer in pitch than final syllables to neighbouring tone 3 syllables.

Another source of pitch variation within a phrase is **boosting**. I will suppose that this is an optional phonological rule which raises the pitch of obligatory tones, or more accurately, increases the interval between tones. For the sake of simplicity we can assume that, like key changing, it does this by raising the top line but not the baseline. The effect of boosting may be illustrated with a phrase carrying the Drop tune. In 22, but not 21, the maximum pitch of the nuclear syllable is considerably higher than the maximum pitch of the prenuclear accented syllable.



I do not treat 21 and 22 as exemplifications of different tunes (or parts thereof), as do several expositions of English intonation. This would be unsatisfactory because the effect of replacing 21 with 22 in a discourse is quite different from the effect of replacing it with a phrase bearing what is clearly a different tonal morpheme. Utterance 22 is appropriate insofar as it is appropriate for the speaker to display excitement, or as Cruttenden (p. 121) has it, "involvement". Replacing the Drop with a Dipper, on the other hand, makes the phrase more appropriate for a question than a statement. A second reason for not treating the boosted and unboosted Drops as different tonal morphemes is that boosting applies to other tunes as well, with the same phonetic and pragmatic effects. The following traces illustrate normal (23) and boosted (24) versions of the Dipper tune.



It would be inconsistent to distinguish between a High Drop and a Low Drop without distinguishing between a High Dipper and a Low Dipper. Presumably other analysts have done so because the boosted Drop is difficult to ignore. It is no doubt more common than other boosted tunes, simply because the Drop is the most common tonal morpheme in the standard English dialects.

4.2. Morpheme

Identifying the sequence of tones linked to a phonological phrase as a morpheme represents the hypothesis that they constitute a minimum grammatical unit. By *grammatical unit* I mean simply a linguistic unit which is not a phonological unit. A minimum *x* is an *x* which contains no lesser *xs*. To claim that an intonational tune is a morpheme is therefore to claim that (1) it is a grammatical unit, and (2) it contains no lesser grammatical units. I will discuss these two claims in sections 4.2.1 and 4.2.2. This hypothesis is by no means new. Trager and Smith explicitly identify their intonation contours as morphemes. Most students of intonation have avoided the issue by using only terms such as *tune*, *tone group*, *intonation contour*. This is not merely evasive. It also implies that linguistic units which consist of tones are *sui generis*-- completely different from those consisting of phonemes. If tonal morphemes were indeed quite different in function from all segmental morphemes, it could be argued that the term *morpheme* is inappropriate. I will suggest in 4.2.3 that they are not so different: that there is a class of segmental morphemes which behaves in much the same way.

4.2.1. Grammatical unit The hypothesis that intonational tunes are grammatical units embodies in its turn two subclaims: (i) that they are linguistic units, and (ii) that they are not phonological units. The first of these is probably the least controversial. Since there has been very little discussion of these questions, linguists' replies are mostly implicit. For example, we can assume that those linguists who have a name for the sequence of tones attached to a phonological phrase or word group recognize it as a unit. Examples of such terms are O'Connor and Arnold's (1973) *tone group*, and Trager and Smith's (1951) *intonation contour*. The position of others is not so clear. Both Halliday (1967) and Crystal (1969a), for example, have a term for the obligatory part of a tonal morpheme-- *tone* and *nuclear tone*, respectively-- but not for the whole sequence of tones attached to a phonological phrase-- *tone group* and *tone-unit*, in their terms. It is reasonable to take this as an indication that they do not recognize this sequence as a unit, but even they do not explicitly deny that it is.

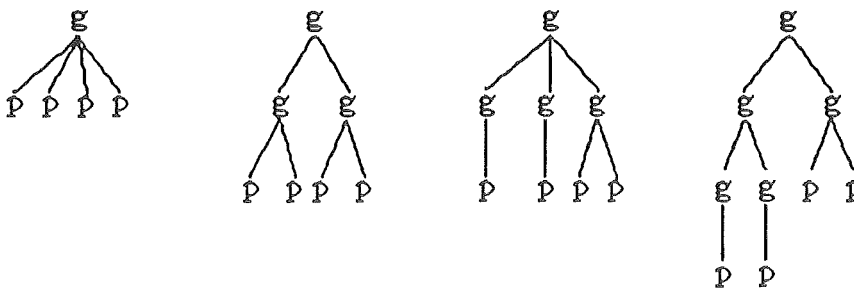
The most obvious advantage of treating tunes as units is that it simplifies the statement of the relationship between intonation and phrasing. Every phonological phrase contains, or is associated with, one and only one tonal morpheme. If we did not recognize a unit corresponding to the tonal morpheme, the relationship would appear rather more complicated, somewhat as follows. Every phonological phrase contains one nucleus (or some such term); some phrases also contain one head; some phrases also contain one prehead. In the absence of evidence to the contrary, we adopt the hypothesis that there is a one-to-one relationship between intonational units and phonological phrases.

It seems that those linguists who take tunes to be linguistic units also take them to be grammatical units. The only evident alternative is that they are phonological units. Such a position runs into difficulties. That tunes have meanings-- in the widest sense, including syntactic, truth-conditional, and pragmatic meaning-- is a commonplace in the study of intonation. This fact is a problem for a theory in which tunes are phonological units, because the hypothesis that phonological units do not have meanings is one of the most fundamental in general linguistic theory. It is grammatical units, such as morphemes and words, that have meanings. The existence of a few morphemes containing only one phoneme poses no threat to this assumption. If this assumption is true, then tunes are not phonological units. Since identifying tunes as phonological units would falsify a well-established and fundamental hypothesis, a very strong case would have to be made.

In fact, the position that tunes are phonological units does not appear to be popular. The inexplicitness of much work on intonation makes it difficult to know whether any linguists have held it. I have presented the argument against it anyway, both for the sake of completeness, and, more importantly, because the same argument applies to a position which has been popular. I refer to theories which identify the obligatory part of a tune as a phonological unit, i.e. as a tone. This claim is rarely made explicitly, but the use of the term *nuclear tone* represents an implicit claim that the part so described is a phonological unit. If this implication is not intended, then the choice of terminology is misleading. More to the point, linguists who refer to nuclear

tones readily refer also to the meanings of such units. None of them appears to recognize that this flies in the face of a fundamental assumption in the study of language. If the hypothesis is true, it necessitates the complication of a simple and wide-ranging hypothesis about linguistic structure. Because of these potential consequences, the proponents of the hypothesis that nuclear tones have meaning need to demonstrate not that these sequences may be treated as phonological units, but that they must be. To my knowledge, they have not even attempted such a demonstration.

4.2.2. Minimum grammatical unit The hypothesis that the intonation contour of a phonological phrase represents a morpheme entails not only that it is a grammatical unit (in the sense defined above), but also that it does not contain any other grammatical unit. Thus we suppose that the tones which constitute a tonal morpheme are its immediate constituents. The left-hand diagram below illustrates the structure of such a tonal morpheme, with *g* standing for *grammatical unit*, and *p* for *phonological unit*. The other three diagrams illustrate other possible structures, all consistent with the hypothesis that tunes are grammatical units, but not with the hypothesis that they are minimum grammatical units.



The assumption that the immediate constituents of tunes are phonological units (tones) rather than grammatical units seems to me a reasonable working hypothesis. To assume otherwise is to postulate at least one more layer of structure. Such a layer needs to be justified. this view of the matter is not widely shared. Both Liberman (1978) and Ladd (1980), for example, take the view that a tune does contain lesser grammatical units. That is, they consider that the immediate constituents of a tune are not phonological units but

morphemes-- or analogous to morphemes, as Liberman (1978:88) has it. It follows from this that the whole tune cannot be a morpheme, as we are using the term. One of these constituents corresponds to the postulated unit traditionally called the *nuclear tone*. It appears, then, that the position taken by Liberman and Ladd is more traditional than mine.

The traditionalness of the alternative view seems to have obscured the fact that it requires support. My position is really a null hypothesis. I do not need to make a case for not hypothesizing an extra layer of structure. The onus is on those linguists who postulate extra morphemes to present the facts which their theory explains better than the simpler theory. A survey of the literature reveals little in the way of serious attempts to make a case for the hypothesis that tunes contain lesser grammatical units. I will deal here with a few arguments, some of them only implicit.

It seems to be widely assumed that the prenuclear tones must be separate from the obligatory tones simply because they are optional. This argument holds no water at all. There is in English a very common morpheme which includes a phoneme which is optional in much the same way: *a(n)*. The second phoneme appears only if there is a vacant consonant slot immediately following the vowel slot occupied by the first phoneme. I do not believe that any linguist has suggested that the two phonemes of *an* belong to different morphemes. The prenuclear tones of a tonal morpheme likewise occur if and only if there are unassociated syllables preceding the nuclear syllable of a phonological phrase. Optional is not the best word for either of these items, but we will not spend time at this point searching for a better term. The point is that even in English we can find an example of a form which is clearly minimal and contains an optional element. (The dropping or adding of the final phoneme of *an* cannot be done by an ordinary phonological rule because it does not occur with other words ending with /n/.)

Cruttenden points out that the obligatory part of a tune makes a greater contribution to its meaning than the optional prenuclear part. This is not surprising, since it is the only part that is always present, but we cannot conclude from this that it constitutes a separate morpheme. To see this, try applying the argument to *an*. If one insists that one of the phonemes is

semantically more important than the other, it must be the /æ/, since the /n/ is frequently absent without damage to the meaning. But this has quite rightly not led anyone to argue that /æ/ and /n/ are separate morphemes.

Another potential argument is based on the relative lack of variety in the optional part of the tune. All of the common English tunes begin in the same way, with an optional 1-tone and an optional accent 3-tone. The fact that these elements are shared by so many tunes suggests to some that they represent the same linguistic form. It follows that each tune consists of more than one linguistic form. The trouble with this argument is that a phonetic/formal common factor is not enough to identify parts of different forms as the same morpheme. There must be a semantic/functional common factor as well. Only three of the tonal morphemes discussed in this study depart from the pattern just mentioned. There is no evidence that all those tunes which begin with (1)(3) have something in common functionally or semantically which sets them apart from those that do not contain these optional tunes.

By the same token, I have not seen a demonstration that particular sets of obligatory tones correspond to particular meanings. The Low Bounce and the Take Off (and perhaps the Contradiction Contour) have the same obligatory tones-- 13-- and no other tune shares this set. Is there a corresponding similarity of meaning or use between these two tunes which is not shared by any other tonal morpheme? Showing that there is would be a difficult task, and no one has undertaken it, as far as I can determine. Referring to these tunes, Cruttenden (1986:57) claims that particular "sentences are gentler and more tentative when said with a rise than when said with a fall". This is a long way from showing that this gentleness is characteristic of all and only sentences spoken with a "low rise". Apparently it is generalisations like the one quoted that Cruttenden is referring to when he says (p. 56)

Almost all intonational analysts agree that . . . an analysis purely in terms of whole tunes fails because it misses important generalisations dependent on the occurrence of similar tones starting from the nucleus.

It is not clear that such generalisations are worth capturing.

I would happily devote more space to this issue, but the lack of arguments from the opposite side makes it difficult to do so. It may well be that the extra

structure is necessary, or that it does not in fact represent a more complicated theory, but it would be unscientific to abandon what seems to be a simpler theory when the superiority of the alternative has not been demonstrated. In this case the traditionalness of the complicated theory may explain why it is held, but it does not justify continuing to hold it.

4.2.3. Stylistic morphemes Many students of intonation avoid the question of whether intonational tunes are morphemes-- or words, or phrases, for that matter-- by using only terms such as *contour*, or *tune*, or *tone group*. Such labels suggest that these entities are functionally as well as formally *sui generis*. It might be argued that it is not appropriate to describe intonational tunes as morphemes, irrespective of their structure, because they function in a quite different way from linguistic units made up of phonemes. I wish to conclude my justification of the term *tonal morpheme* by disposing of this argument.

It is certainly true that tonal morphemes are functionally unlike the most familiar kinds of grammatical units. For example, replacing one tonal morpheme with another does not appear to affect the grammaticality of an associated syntactic unit. Tonal morphemes therefore differ from inflexional morphemes such as the suffix on *knows*. Removing this morpheme or replacing it with another, such as *-ed*, results in an ungrammatical sequence.

The elephant knows the answer.

*The elephant know the answer.

*The elephant knowed the answer.

Neither does replacing one tonal morpheme with another change the semantic interpretation of an associated sentence or phrase, assuming a truth-conditional account of semantic interpretation. That is, the intonation of a sentence does not affect the conditions under which it is true. By contrast, replacing *elephant* with *tiger*, or *know* with *want* does change the truth conditional meaning of a sentence. In other words, tonal morphemes are not semantically interpreted, and in this they differ from most English lexemes. This explains why it is so difficult to identify the meaning of tonal morphemes. They do not contribute to the propositional core of utterances, i.e. to "what is said". Rather they indicate to the hearer what implicatures the speaker

intends the hearer to draw from what is said. Thus they indicate whether an utterance is to be taken as an objection, a question, a request for repetition, a threat, or simply an assertion. Let us say that tonal morphemes are pragmatically interpreted rather than semantically interpreted.

This characteristic is not, however, peculiar to linguistic units which consist of tones. There are also "normal" lexemes-- lexemes consisting of phonemes-- which are not semantically interpreted, but pragmatically interpreted. They are relevant to the interpretation of utterances, i.e. sentences in contexts. They contribute not to the truth conditions of sentences, but to the appropriateness conditions of utterances. Consider, for example, one of the words pronounced /ðen/-- call it *then*₃.

John has an elephant, then.

John has an elephant.

Assuming a unique referent for *John*, removing *then*₃ has no effect on the semantic interpretation. That is, there are no conditions under which one of these sentences is true and the other false. There will, however, be conditions under which one is appropriate and the other is not. Whether a sentence with *then*₃ is acceptable depends on pragmatic factors, i.e. features of the situation in which it is used. There appear to be many such items in the English lexicon, but as a group they have received less attention than comparable items in other languages-- in particular, the particles of German and ancient Greek. Schubiger (1965, 1980) has shown that changing the tune of an English sentence often has the same effect as changing (or adding or deleting) a modal particle in the corresponding German sentence. This is what we should expect if both English tonal morphemes and German modal particles are pragmatically interpreted.

We may call these items stylistic morphemes, because they are interpreted in the same way as what Rochemont (1978), Kuiper (1982), and others call stylistic rules. Examples of such rules are Heavy NP Shift, PP Extraposition from NP, and Modifier Repetition. One of the characteristics of such rules is that they do not affect grammaticality or semantic interpretation. Whether their output is acceptable depends on pragmatic factors.

I suggested above that some of the terms used instead of *tonal morpheme* implicate that intonational units are not only formally distinct from nontonal

grammatical units, but functionally different, as well. We have seen that such an implicature is misleading. English tonal morphemes are pragmatically interpreted, but so are many lexemes which consist of phonemes. I conclude that it is fitting to use the same terminology-- *morpheme*, *word*, *phrase*, *lexeme*, for example-- for forms composed of tones as we do for forms composed of phonemes.

I began this chapter by introducing the basic principles of autosegmental phonology, in which a phonological representation consists of several autonomous sequences of elements, and outlining tone association rules which link the elements of different sequences to one another. In the next section the features which I use to represent the prominence of syllables in phonological phrases were presented. I then discussed in some detail the representation of three English tonal morphemes, showing how my proposals about the form of these morphemes and assumptions about the operation of the tone association rules make accurate predictions about the major pitch variations in phrases spoken with these tonal morphemes. In the course of this exposition, the terms *nucleus tone*, *accent tone*, and *end tone* were explained. This was followed by a discussion of the phonology and phonetics of tone, with some attention to causes of variation in the pitch of particular tones. Finally, I argued that the term tonal morpheme is an appropriate one for these linguistic entities.

Chapter 2

Variable Stress

Words such as *unknown*, *archbishop*, *afternoon*, and *absolutely* are traditionally said to exhibit stress shift, or variable stress. These terms refer to the fact that in some utterances of these words the second or third syllable is the most prominent, while in other utterances the first syllable is the most prominent. In this chapter I will show how the theory adopted in this work provides an explanation for this phenomenon which is superior to other treatments of it. My primary purpose is to provide some support for the type of prosodic representation adopted in Chapter 1 as the input to the tone association rules.

We begin, in section 1, with a description of the phenomenon. The next section presents a set of rules to account for the data outlined in the first one. In section 3 the inadequacies of three competing explanations are demonstrated. In section 4 I consider some of the inadequacies of our theory, and modify it so as to eliminate them.

1. The Problem

I will begin by outlining what seem to be the crucial data which our theory will need to account for. In Chapter 1, I introduced a set of three binary features for characterizing prominence in phonological phrases: [nucleus], [accent], and [strength]. I will refer to these as **phrasal prominence features**. Since they are hierarchically related, they distinguish four degrees of prominence, each of which is exemplified once in a citation utterance of the phrase *afternoon tea*.

	Afternoon tea			
Nucleus	-	-	-	+
Accent	+	-	-	+
Strength	+	-	+	+

A transcription system was introduced, to save space. The following representation is equivalent to the one immediately above.

'After'noon 'tea | |

I use the terms **nuclear**, **accented**, and **strong** as equivalents of [+nucleus], [+accent], and [+strength], respectively.

Consider now the word *confidential*. A citation utterance of this word would be represented as follows. I will omit the indication of the phonological phrase boundary in these transcriptions, because we will deal only with single phrase utterances.

1. 'confi'dential

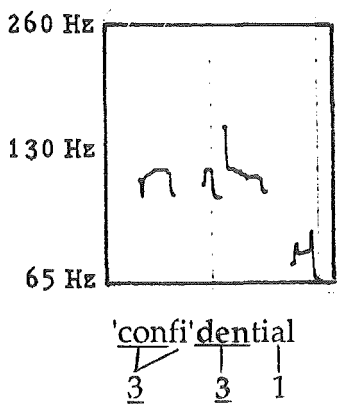
The first syllable is accented but not nuclear. The third syllable is both accented and nuclear. It follows from our definition of these terms that the first syllable is more prominent than the second, which is unaccented, and less prominent than the third. It is not claimed that experimentally obtained assessments by native speakers of the relative prominence of uttered syllables are determined solely by the specifications of phrasal prominence features. Phonological prominence is an important-- possibly the most important-- determinant of native speakers' reported judgements of the relative prominence or loudness of particular syllables in particular utterances, but it is not the only one. It is clear, for example, that word boundaries play a part. There is less agreement about the relative prominence of syllables when they are in different words than when they are in the same words. It seems likely that the number of intervening syllables has a similar effect. In fact, there are presumably a great many non-phonological factors which may affect people's judgements of the prominence of a particular syllable in a particular utterance. I will not attempt to provide experimental support for these hypotheses; they do not appear to be controversial. I wish simply to make it clear that phonological prominence is not to be identified with particular judgements of prominence, and to suggest that imperfect correspondence between them is readily explicable.

Because the features [accent] and [nucleus] are defined not only in terms of

prominence, but also by their role in the distribution of tonal elements, claims about the distribution of these features are easier to test than in a theory which has no equivalent of the feature [accent]. For example, the representation above predicts a particular distribution of tones if the phonological phrase *confidential* is spoken with the tonal morpheme typical of citation utterances, the Drop: (1)(3)31. The initial tone association rules will link the nucleus tone to *den*, because that syllable is nuclear. They will also link the first 3-tone to *con*, because it is accented. The wellformedness rules of tone association will attach *fi* to the first 3-tone, and *tial*, to the final 1-tone. The initial 1-tone remains unattached.

1a. 'confi'dential
 3 3 1

If we say that syllables attached to a 3-tone are high, we predict that *con* will be the first high syllable, and *den* the last high syllable in the citation form of *confidential*. This is indeed what we find.



It is agreed, then, that in a citation utterance of *confidential*, the third syllable is the most prominent, and hence more prominent than the first syllable. The third syllable is also most prominent in a citation utterance of *it's very confidential*:

2. it's 'very 'confi'dential

In this transcription the first syllable of *confidential* is represented as strong but not accented. This amounts to a claim that it is not only less prominent than *den* but also less prominent than the first syllable of *very*. If this phrase has the Drop tune, *ve* will be the first high syllable, and *den* the last high

syllable.

2a. it's 'very 'confi'dential

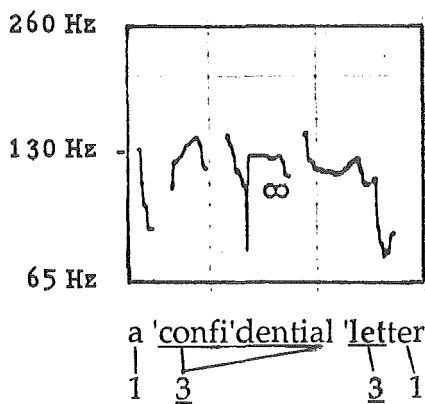
1 3 3 1

We may find that native speaker judgements of the relative prominence of *ve* and *con* (or *ve* and *den*), are less uniform than judgements on the relative prominence of *con* and *den*. There is a simple explanation for this sort of variation. It is easier to assess the relative prominence of two syllables if they are in the same word than if they are not.

In other phrases containing the word *confidential*, the first syllable appears to be more prominent than the third. This happens when the first syllable is accented and the third is not, as in the citation form of *a confidential letter*.

3. a 'confi'dential 'letter

Spoken with a Drop tune, the first high syllable will be *con*, and the last high syllable will be *let*.



Finally, there are utterances of *confidential* in which the first and third syllables are equally prominent. This occurs when neither of the syllables is accented, which happens when the word precedes the nuclear syllable of the phrase, as in 4, and when it is entirely postnuclear, as in 5.

4. 'seven 'confi'dential 'letters

5. it's 'very 'confi'dential

We shall see that the linguists discussed below either do not make this observation, or having made it, do not perceive its relevance to the problem. The failure of other theories to account for such data is a serious defect. Since

this claim is so crucial, I will present a simple way of testing it later in the chapter.

2. An Explanation

One might attempt to account for the phenomenon described in section 1 by supposing that speakers of English learn a set of variant pronunciations for every such word, and perhaps for all words. That is, the lexical entry for each such word would contain representations of several prominence patterns. For each pattern the lexical entry would also have to contain information about the contexts in which it occurs, because the patterns are not in free variation. This sort of explanation adds a great deal to the amount of phonological information in a lexical entry, information which native speakers must commit to memory.

This sort of explanation is counterintuitive as well as uneconomical. Our intuition is that each word has one pronunciation, not several. There are exceptions to this, of course, but they serve only to underline the inadequacy of the approach described above. There are words which are perceived to have more than one stress pattern, words such as *controversy*, *contribute*, *mustache*, and so on. We want our theory to distinguish between words like *controversy*, which cause arguments amongst native speakers and provoke letters to the editor, and words like *confidential*, which do not. Native speakers notice the different pronunciations of *controversy*, as they notice the segmentally different variants of *either*, *dynasty*, and *economical*. We account for this by providing such words with two phonological representations in the lexicon. For the vast majority of words, including *confidential*, we can postulate a single phonological form in the lexicon.

The explanation which I will propose accounts for a relatively complex phenomenon in terms of a simple set of rules combined with simple phonological representations in the lexicon. I take it as given that all words vary in pronunciation-- both prosodically and segmentally-- from utterance to utterance, and that a great deal of this variation is systematic. It is now traditional to characterize the pronunciation, or phonetic form, of a word in a

particular utterance as the result of various phonological rules operating on the underlying phonological form of that word. These phonological rules are not actually operations, but rather statements of correspondences between representations of the same linguistic entity at different levels of description. The phonological rules which account for the variations which concern us here can be called **accentuation rules**. These rules assign the phrasal prominence features which characterize the syllables of phonological phrases. Since the tone association rules refer to the features [accent] and [nucleus], it is clear that they must follow the accentuation rules. That is, the representations at the level of input to the tone association rules are less abstract, more fully specified, more phonetic, than the representations at the level of input to the accentuation rules.

There are two kinds of accentuation rule. In subsection 2.2 I will discuss the Nuclear Accentuation Rule, which makes certain syllables [+nucleus] (and, redundantly, [+accent] and [+strength]). The Prenuclear Accentuation Rule, which makes certain syllables [+accent, +strength], is dealt with in 2.3. Subsection 2.4 demonstrates the application of the rules to the problem of stress shift. Before I take up these rules, it is necessary to introduce what I call lexical prominence features, since the accentuation rules refer to these features. In other words, the syllables of representations at the level of input to the accentuation rules are specified for these features.

2.1. Lexical Prominence Features

These are called lexical prominence features because they characterize the syllables of lexemes (or rather, those lexemes that consist of phonemes). The lexical entry for each lexeme contains phonological, syntactic, semantic, and pragmatic information. Included amongst the phonological information will be, for each syllable, its specification for the two features [stress] and [primary]. Let us suppose that each syllable is either [+stress] or [-stress]. Let us further suppose that one, and only one, of the [+stress] syllables in each word is also [+primary]. All the other syllables will be [-primary]. All [-stress] syllables are redundantly [-primary]. For monosyllables, the specifications of both features are redundant. For lexemes with only one [+stress] syllable, the specification of

the feature [primary] is redundant. I show below how a selection of English words would be specified for these features.

PRIMARY	+	-	+	-	+	-	-	+	-	-	+	-	-
STRESS	+	-	+	+	+	-	-	+	+	-	+	-	+
	bacon		moron		gentleman			newspaper		superman			

PRIMARY	-	+	-	+	-	+	-	-	+	-	-	+	
STRESS	-	+	+	+	-	+	-	+	+	-	+	-	+
	refuse(V)		re-fuse		important			archbishop		afternoon			

As Chomsky and Halle (1968) show, it is possible to predict the distribution of these feature specifications for a large number of words, given information about their phonological structure and syntactic category. Since they are not all predictable, I will suppose that the syllables of every lexeme are specified for these two lexical prominence features, and that the correspondences just mentioned are embodied in lexical redundancy rules in the sense of Jackendoff (1975).

2.2. Nuclear Accentuation Rule

The accentuation rules require information about the location of phonological phrase boundaries, the lexical prominence feature specifications, and which constituents are focussed. There are many languages for which it is necessary to postulate a syntactic feature like [focus] to account for the distribution of what are called focus particles or affixes. As Comrie (1981:260) points out with reference to the Siberian language Yukagir, the distinction between different focusses is most obvious in question and answer sequences. If a statement is spoken in response to a question, the focussed elements will be those conveying information not in the question. Thus *ran away* is focussed in 6b below, and a focus prefix /me/ is attached to the word for *ran away* in the Yukagir translation; but the focus affix /leŋ/ goes with the word for *deer* in 7b, since that word carries the new information in that context. Unfortunately, Comrie does not supply the Yukagir questions which I label 6a and 7a.

6a /	/	7a /	/
"What did the deer do?"		"What ran away?"	
6b / ileŋ me-kötege -j /		7b / ile-leŋ kötege -l /	
deer FOC run-away 3SG		deer FOC run-away 3SG	
"The deer ran away."		"The deer ran away."	

(The form of the affix depends on the syntactic structure. The verb suffixes which indicate person and number agreement also vary according to which constituent is focussed, as these examples show.)

Let us suppose that [focus] is a syntactic feature which determines the location of focus affixes, and also functions crucially in defining the appropriateness of a particular utterance in a given context. The discourse fragments 6 and 7 are well-formed. Both 6b and 7b are grammatical, but if 6b were replaced by 7b following 6a, the resulting discourse would be ill-formed.

In English, too, the appropriateness of an utterance depends partly on which constituents are focussed. English does not, it seems, have focus particles, but [focus] does determine the location of the nuclear accent. Thus if a verb phrase which is present in the question is repeated with a nuclear syllable in the response, the response is inappropriate in that context. Thus the sequence 8 is pragmatically well-formed, while 9 and 10 are ill-formed, because *ran away*, present in the question, contains a nuclear syllable in the answer.

8. What ran away? The **deer** ran away.
9. What ran away? The deer ran away.
10. What ran away? The deer ran away.

I have said that in English the location of the focussed constituents determines the location of the nuclear accent, but I have not yet made the relationship clear. We have already seen that the same sentence-- in one sense of the word "sentence"-- can be an appropriate reply to different questions. Consider the following question and answer pairs. In each answer I have italicised the words which are not in the question, and can therefore be taken to represent new information, and to be focussed. Boldface indicates the syllable which would be nuclear in an appropriate response to each question.

11. Who dropped the camera? *Dave* dropped the camera.

- | | |
|-------------------------------------|----------------------------------|
| 11. Who dropped the camera? | <i>Dave</i> dropped the camera. |
| 12. What did Dave do to the camera? | Dave <i>dropped</i> the camera. |
| 13. What happened to the camera? | Dave <i>dropped</i> the camera. |
| 14. What did Dave drop? | Dave dropped <i>the camera</i> . |
| 15. What did Dave do? | Dave <i>dropped the camera</i> . |
| 16. What happened? | Dave <i>dropped the camera</i> . |

It is clear that the nuclear syllable is consistently in the last focussed word. If we consider the last three pairs, we see that an utterance of *Dave dropped the camera* in which the last word contains the only nuclear syllable corresponds to three sentences (at least): one in which the final noun phrase is focussed, one in which the verb phrase is focussed, and one in which the entire sentence is focussed. Such an utterance would be an appropriate reply to 14, 15, or 16, but not to 11, 12, or 13. We can account for this by postulating a rule which assigns a nuclear accent to the last [+focus] word in a phonological phrase. The syllable of the selected word which receives the specification [+nucleus] is the [+primary] one. Thus the nuclear accentuation rule takes the following form.

Nuclear Accentuation Rule

Make the [+primary] syllable of the last [+focus] word in a phonological phrase [+nucleus].

This formulation skates over a great many complications, and is only testable when slotted into a relatively complete theory of grammar operating in conjunction with a theory of communication, including discourse structure, pragmatics, and so on. To pursue them would take us far beyond the scope of this work, since the literature on focus and related topics is enormous. This rule is perfectly adequate for our purposes in this chapter. I will mention a couple of these complications in order to obviate misunderstanding.

In the preceding examples, it is possible to identify focussed elements with those representing what is often called new information, the latter being information which is not old, or given, or already present in the discourse. In fact it is not possible to maintain this identification of focus and new information. On the one hand, we find new information which is not focussed. The following utterance would be appropriate in a discourse in which the phrase *this morning* represented new information, but the fact that

which the phrase *this morning* represented new information, but the fact that it follows the accented syllable indicates that it is not [+focus].

17. I ran into John this morning.

We conclude that focus does not include all the new information. On the other hand, it is clear that focus does not include only new information, in the transparent sense of the term. We find examples of focussed elements which do not represent new information, such as *she* in 18.

18. John hit Mary, and then she hit him.

It does not follow from this lack of correspondence between focus and new information that there are no systematic correlations between them. There do indeed appear to be systematic correlations, but the assumption that focus has a unitary semantic or pragmatic definition must be abandoned. I will not pursue this matter further here, because the resolution of such problems does not affect the application of the nuclear accentuation rule in the examples which concern us. I will make the uncontroversial assumption that all the material in a citation utterance is focussed. The same assumption appears to lie behind most discussions of stress shift, but few linguists make it explicit.

Secondly, it is clear that the nuclear accentuation rule above will only make one of the syllables in a phonological phrase nuclear if the phrase contains [+focus] material. It is possible that, as Rochemont (1986) suggests, focussed constituents can be identified by structural means as well as by the feature [focus]. In particular he suggests that *wh*- questions need not contain [+focus] material, because the *wh*- phrase is focussed by virtue of its syntax. If this is correct, then the nuclear accentuation rule will need to be supplemented or amended. For the purposes of this discussion, however, we can stay with the simple version, and avoid those structures which Rochemont identifies as syntactically focussed. In short, we will continue to assume that every nuclear syllable identifies a [+focus] expression.

2.3. Prenuclear Accentuation Rule

Let us turn now to the rule which assigns nonnuclear, or prenuclear, accents. In the first place, it is clear that the prenuclear accented syllable can precede the focussed part of a phonological phrase. A word may be accented

precede the focussed part of a phonological phrase. A word may be accented even though it is not focussed. For example, in the following discourse it is inappropriate for *Dave* to be [+focus] in the reply, but not inappropriate for it to be accented. The discourse appears to be well-formed.

19 a. 'What did 'Dave 'drop?

b. 'Dave 'dropped the 'camera.

As an initial hypothesis we might suppose that the first word of the phrase is accented. We can show that this is incorrect by replacing *Dave* with a common noun phrase, such as *the agent*, or *my husband*. If only *the camera* is in focus, the prenuclear accent will be on the second word. Let us therefore suppose that the prenuclear accent is assigned to the first content word, and in particular, to the first stressed syllable of that word.

Prenuclear Accentuation Rule

Make the first [+stress] syllable of the first content word in the phonological phrase [+accent].

The term *content word* here includes noun, verb, adjective, and adverb, but not, apparently, preposition. An utterance like 20, with *in* unaccented, would be appropriate whether the preposition was focussed or not.

20. in the 'river

This rule is inadequate, however, because it cannot account for phrases in which the first accented syllable precedes the first content word. It predicts that words which are not content words are never accented. This is clearly incorrect, because phrases like 21 where a function word is accented, do occur.

21. 'in the 'river

Let us suppose that an initial function word can be accented, provided it is [+focus]. We can accomplish this by splitting the rule into two parts.

Prenuclear Accentuation Rule II

(i) If the first [+focus] word precedes the first content word, make the first [+stress] syllable of the first [+focus] word [+accent, +strength].

(ii) If the first [+focus] word does not precede the first content word, make the first [+stress] syllable of the first content word [+accent, +strength].

Thus we predict that 22, with *on* accented, is an appropriate response to 23, because *on* is appropriately focussed in such context. We further predict that 22 is inappropriate as a response to 24, because *on* is not an appropriately

focussed element in that context. This appears to be correct.

22. 'no | | 'on the 'trunk | |

23. Are the scissors in the drawer?

24. Are the scissors on the table?

Finally, we will need a rule which replaces the specification [+stress] with the specification [+strength] in those syllables which have not been accented.

Let us suppose that the following rule applies after the accentuation rules.

Strength Rule

Make all [-accent, +stress] syllables [+strength].

Later in this chapter I shall propose modifications which will make this rule unnecessary, but let it stand here for the sake of completeness. We will also need a rule capable of making strong syllables weak, to cope with phrases like 25 where *young*, which must be [+stress], is weak.

25. There 'was a young 'fellow called 'Hyde

I do not intend to discuss this sort of rhythm rule, except to point out that an additional optional rule is appropriate because such phrases are recognized as exceptional.

2.4. Application to problem

Stress shift will appear to occur in words which have at least one [+stress] syllable to the left of the [+primary] one. I list below some of the words which have been used in the literature to illustrate this phenomenon.

Primary	-	+	-	+	-	-	-	+	-	-	+	-
Stress	+	+	+	+	-	+	-	+	+	-	+	-
	unknown		archbishop			afternoon		absolutely				
	thirteen		Westminster			clarinet		continental				
	fullgrown					Tennessee		achromatic				
	Chinese							telegraphic				

There are many longer words which meet the criterion above, but they are rarely discussed. I give a few five-syllable and seven-syllable words here.

P	- - - + -	- - - + -	- - - - + - -	- - - + - - - -
S	+ - - + -	- + - + -	+ - + - + - -	+ - - + - - + -
	magnification	affiliation	unreliability	antiamericanism
	psycholinguistic	apotheosis	autobiographical	

I have used the word *confidential*, which is like *absolutely* and so on, in my illustrations. Comparable examples can be constructed for all the words listed, and we shall see some in subsequent sections. At this point I will repeat examples 1 to 5 in order to confirm that the features and rules postulated above do make the predictions which we require.

1. 'confi'dential
2. it's 'very 'confi'dential
3. a 'confi'dential 'letter
4. 'seven 'confi'dential 'letters
5. it's 'very 'confi'dential

I will make the reasonable assumption that none of the function words is focussed, and all of the content words, with the exception of *confidential* in 5. Other investigators adopt similar assumptions about citation utterances, but do not always make them explicit.

In 1 and 2, *confidential* is the last focussed word. The Nuclear Accentuation Rule makes the third syllable [+nucleus] because it is [+primary]. In 1, *confidential* is also the first content word, so the Prenuclear Accentuation Rule makes the first syllable [+accent]. In 2, however, the first content word is *very*, so the first [+stress] syllable of *very* is accented by the PAR. In 3, *confidential* is the first content word again, but not the last focussed word. The [+primary] syllable, the third, therefore remains unaccented, while the first is accented because it is the first [+stress] syllable of the first content word. It is thus more prominent than *den* but less prominent than the first syllable of *letter*, which the NAR makes [+nucleus].

In both 4 and 5, *confidential* is neither the first content word nor the last focussed word. For this reason neither of its [+stress] syllables is accented by the accentuation rules, and the Strength Rule makes them both [+strength]. The first syllable of *letters* in 4 is [+nucleus], because it is the [+primary]

syllable of the last focussed word. The PAR accents the first [+stress] syllable of *seven*, the first content word. This syllable is therefore more prominent than the strong syllables of *confidential*, but less prominent than *let*. *Very* in 5 is both the last focussed word and the first content word. Since the first syllable is [+primary], the NAR makes it nuclear. Since the same syllable is the first [+stress] syllable, the PAR applies vacuously.

The claim that the two strong syllables of a double-stressed word are equally prominent in phrases like 4 and 5 can be tested with pairs of words which differ prosodically only in the location of the [+primary] syllable. There are many disyllabic noun-verb pairs in which the noun has the first syllable primary and the verb has the second syllable primary. Most of these are unsuitable for the test in one way or another. Commonly the first syllable of the verb may be weak, for instance. Often the verb may be stressed either way, and so on. Fortunately, we can avoid these difficulties, because there is at least one pair of words which belong to the same lexical category, and are distinguished only by the location of the feature specification [+primary]. They are the compound nouns *party line* and *party line*. A *party line*₁ -- with the first syllable primary -- is a telephone line serving more than one subscriber. A *party line*₂ -- with the last syllable primary -- is the set of essential beliefs or policies of a political party. My theory predicts that these words will be homophonous whenever they do not contain a nuclear syllable. It predicts, then, that the following pairs of phrases will be homophonous.

- | | |
|---|---|
| 26a. The <u>party</u> line ₁ is <u>dead</u> . | b. The <u>party</u> line ₂ is <u>dead</u> . |
| 27a. He's <u>off</u> the party line ₁ at <u>last</u> . | b. He's <u>off</u> the party line ₂ at <u>last</u> . |
| 28a. We <u>haven't</u> <u>got</u> a party line ₁ . | b. We <u>haven't</u> <u>got</u> a party line ₂ . |

Informal testing confirms that this prediction is correct. Native speakers of English are not able to distinguish an utterance of 26a (27a, 28a) from an utterance of 26b (27b, 28b). This is not to say that it is impossible to say such sentences in a way which will disambiguate them. It is likely that most pairs of homophones can be disambiguated by some means, such as stretching or pausing, but this is irrelevant to the correct representation in terms of phrasal prominence features. Both 26a and 26b are manifested as 26, and so on. It

would be a simple matter to test this prediction formally.

26. The 'party 'line is 'dead

27. He's 'off the 'party 'line at 'last

28. We 'haven't 'got a 'party 'line

I must emphasize that the rules we have tested were not devised to account for variable stress. They apply to all phonological phrases, whether they contain "variable stress" words or not. The rules will accent the first [+stress] syllable of the first content word (assuming no [+focus] word precedes it), and this syllable will be phonologically more prominent than any succeeding syllable in the same phrase which is not accented. Other things being equal, this difference is more obvious to the native speaker when the syllables in question belong to the same word. Thus we expect less agreement about the relative prominence of the syllables *con* and *den* in 29 than in 3.

3. a 'confi'dential 'letter

29. O' Connor's 'dented 'Hillman

In a theory where the distribution of tones depends on the location of accented syllables, the claim that a particular syllable is accented is more readily tested than in a framework where [accent] is defined only in terms of prominence.

I mention the generality of the accentuation rules because this fact distinguishes our explanation of variable stress from the others that I will discuss. Other treatments of the phenomenon are ad hoc.

3. Other explanations

Let us now compare the preceding treatment of so-called variable stress with three alternative accounts of the same phenomenon, those of Gimson (1970), Vanderslice and Ladefoged (1972), and Liberman and Prince (1977). These three influential works represent a wide range of approaches to the study of language. Liberman and Prince's work is firmly rooted in generative linguistics. It is based on Liberman 1978, a dissertation supervised at MIT by Chomsky and Halle. Gimson's standard text is in the tradition of Daniel Jones. Gimson himself was a student of Jones and occupied the same chair as his master. The approach of Vanderslice and Ladefoged (1972) is a kind of union

of generative and non-generative, American and British, traditions. I will summarize each account and show that each makes incorrect, or no, predictions where our theory makes correct ones. We begin with the earliest of the three, Gimson's, and move on to the most recent, which is also the most formal. The discussion of this phenomenon in Gimson 1970 is repeated unchanged in the 1980 edition.

3.1. Gimson

Gimson (1970:289) illustrates what he calls "variation of the word's accentual pattern" with the following examples.

- | | |
|--------------------|--------------------------|
| 31a. 'thir`teen | b. 'thir·teen `shillings |
| 32a. 'West`minster | b. 'West·minster `Abbey |
| 33a. 'full`grown | b. a 'full·grown `man |
| 34a. 'after`noon | b. 'after·noon `tea |

Since my three features of phrasal prominence are essentially a formalization of Gimson's features, these transcriptions are equivalent to my representations of citation utterances of these phrases, except that Gimson's grave accent indicates a falling nuclear tone as well as the location of the nuclear accent. The following representations would therefore match 34a, b.



Gimson and I are therefore in substantial agreement about the prosodic representation of the above phrases. I am less happy with Gimson's explanation of the stress shift (p.289).

When a word . . . pattern consists in isolation of a primary accent preceded by a secondary accent, . . . the primary accent may be thrown back to the syllable carrying the secondary stress in isolation, if, in connected speech, a strong accent follows closely.

First of all, Gimson's casual terminological practice does not help to clarify the picture. His use of the term "primary accent" is confusing. In the first clause it seems to replace the term "nuclear accent", as it does elsewhere (e.g. p. 267), referring to something characterizing the most prominent syllable in a phonological phrase. In the principal clause, though, the primary accent is

evidently something marking the most prominent syllable of an uttered word, whether it is nuclear or not. Both of these uses are at odds with his earlier definition (p. 224) of primary accent as a feature indicating the syllable of a word which "will always be associated with a pitch change when the word is said in isolation". In this sense the primary accent is equivalent to our [+primary], in that it belongs to one of the syllables of a lexeme. There is also a problem with the terms "secondary accent" and "secondary stress" in the quotation above; they seem to refer to exactly the same thing. Also potentially confusing is the term "strong accent". Gimson does not explain what he means by this, but in the examples which follow his explanation, all the key words are followed immediately by a nuclear syllable.

Gimson's claim about the cause of the shift can be broken down into two conditional statements:

- 35 i. If the shift occurs, then a nuclear syllable follows closely;
 ii. If a nuclear syllable follows closely, then the shift occurs.

Let us take up each of these statements in turn.

Gimson does not explicitly claim that the stress shift may occur only if a strong accent follows closely, but the lack of any examples to suggest otherwise licenses this interpretation. Certainly he does not mention any other environments which condition the shift. Neither does he make it clear what "closely" means. Let us suppose that it means not more than one syllable away from the stress to be shifted. In these terms, 35i is clearly false. The shift does occur without a closely following nuclear accent, as in 36.

- 36 a. the 'clari'net is 'very 'hard to 'play
 b. 'thir'teen and a 'half is 'excellent
 c. 'abso'lutely ex'pressionless
 d. 'phono'logical re'search

Note, too, that the first subclaim cannot be salvaged by replacing "nuclear" with [+strength]. In 36b and 36c there are two syllables between the shiftable one and the next strong syllable, and in 36d there are three.

Moving on to 35ii, we note once again that Gimson does not explicitly make this claim. That is, he does not say that the shift must occur if a nuclear accent follows closely, but that "the primary accent *may* be thrown back"

under these circumstances. Even this relatively weak claim is false. The stress shift is not optional, given a closely following nuclear syllable. If there is an accented syllable before the word in question, and no intervening phrase boundary, then the stress shift does not occur, even if there is a following nuclear accent. That is, the first syllable of *afternoon* and *magnification* is not the most prominent in the word in phrases like the following.

37. Friday afternoon's fine

38. The degree of magnification is excellent

4. Seven confidential letters

Gimson's explanation of the phenomenon of variable accent is a traditional one. Jespersen (1928:158), in his discussion of the phenomenon, shows that this sort of explanation goes back at least to 1765, when it appears in James Elphinston's *Principles of English Grammar*. Like most traditional accounts, this one is insightful but, in the end, inadequate. Let us move on to an account which is no less traditional, but slightly more formal.

3.2. Vanderslice and Ladefoged

In Vanderslice and Ladefoged 1972, as in the present work, syllables are characterized by two binary features at the level of input to the phonological rules. Vanderslice and Ladefoged call these features [HEAVY] and [ACCENT]. (I capitalize these features in order to distinguish them clearly from mine.) Thus, before their sentence accent rules come into effect, each syllable is either [-HEAVY] or [+HEAVY]. Every [-HEAVY] syllable is also [-ACCENT]. Every [+HEAVY] syllable is either [-ACCENT] or [+ACCENT]. Thus all [+ACCENT] syllables are by definition also [+HEAVY]. At this level of description-- after the word accent rules and before the sentence accent rules-- their [HEAVY] can be identified with our feature [stress]. Their [ACCENT] cannot, however, be identified with our [primary], because a word may contain more than one [+ACCENT] syllable, whereas I have postulated only one primary syllable per word. The [+primary] syllable is always the last of Vanderslice and Ladefoged's [+ACCENT] syllables. The following examples illustrate the difference.

ACCENT	+	-	+	+	+	-	-	+	+	-	+	-	-	+	-	+
HEAVY	+	+	+	+	+	+	-	+	+	-	+	-	+	+	-	+
primary	+	-	-	+	+	-	-	-	+	-	+	-	-	-	-	+
stress	+	+	+	+	+	+	-	+	+	-	+	-	+	+	-	+
	moron	re-fuse	newspaper	archbishop	superman	afternoon										

In section 4 I will show why my analysis is preferable to that of Vanderslice and Ladefoged. I will not pursue the matter at this point, because the inadequacy of their explanation of variable stress can be demonstrated independently.

Vanderslice and Ladefoged propose to account for stress shift by means of a rhythm rule which operates at a point in the derivation after their sentence accent rule. The latter is actually a de-accentuation rule. It changes the specification of the feature [Accent] from + to - in syllables which belong to words which are marked [+mentioned]. It will do no harm to assume that such words would be [-focus] in our description. After this rule applies, the rightmost of the remaining [+ACCENT] syllables is made [+Intonation], i.e. [+nucleus], in my terms. Vanderslice and Ladefoged (1972:827) therefore claim that all postnuclear syllables are [-ACCENT]. Those syllables which are [+ACCENT, +HEAVY] in their underlying form will be [-ACCENT, +HEAVY] in postnuclear position. Thus they predict that the first and third syllables of words like *clarinet*, *Tennessee*, and *telegraphic* will be equally prominent when they follow the nuclear syllable of a phonological phrase. This prediction appears to be correct. I have already discussed what seems to me convincing evidence.

The rhythm rule of Vanderslice and Ladefoged applies after deaccentuation, and, like it, makes [+ACCENT] syllables [-ACCENT]. They motivate the rule in this way. (I underline those syllables which they describe as containing accented vowels.)

It has been a commonplace for upward of two centuries that . . . of three close[d]-spaced heavy syllables, all potentially accentable, the middle one is liable to be de-accented, especially at conversational speed. . . . A phrase which in slow, emphatic speech might be accented like a will be rendered in

rapid speech as b:

a. big bad wolf

+a +a +a

b. big bad wolf

+a +h +a

This also works with polysyllables, causing them to exhibit what is traditionally called "accent recession". Examples are:

A clarinet solo; He plays the clarinet.

Tennessee Williams; The state of Tennessee.

A telegraphic style; His style was telegraphic.

Vanderslice and Ladefoged do not give any more information about their rhythm rule. They do not, for example, make it clear what they mean by "close-spaced". We have already seen that the shift occurs even when there are three weak syllables following the de-accented syllable, as in *phonological research*. More significantly, Vanderslice and Ladefoged's account, referring as it does specifically to three close-spaced accentable syllables, leaves us wondering what happens when there are four such syllables. Does the rhythm rule not operate then? If it does, which triplet of accented syllables does it apply to. There seem to be three possibilities:

- i. The rhythm rule does not operate.
- ii. The rhythm rule applies to the first three [+ACCENT] syllables.
- iii. The rhythm rule applies to the last three [+ACCENT] syllables.

With regard to the first alternative, consider phrases like 39, which would have four [+ACCENT] syllables before the rhythm rule could apply.

39 a. a telegraphic style of speech

b. telegraphing Tennessee

c. unreliability hurts

The feature [+Intonation] will make the final syllable more prominent than the rest. If the rhythm rule does not apply to sequences of four accented syllables, the first and third syllables of *telegraphic*, *telegraphing*, and *unreliability* will remain [+ACCENT]. This makes the prediction that they will be equally prominent in such a phrase, which is patently incorrect.

Consider now the second and third alternatives. A rhythm rule operating on the first triplet of [+ACCENT] syllables can deaccent only the second of these-- *te* in 40. If the rhythm rule operates on the last three of the four, it can deaccent only the penultimate of the four-- *gra* in 40.

40. seven telegraphic messages

Neither formulation will deaccent both *te* and *gra*, i.e. the middle two of a sequence of four [+ACCENT, +HEAVY] syllables. We cannot have it both ways, because the first rule to apply would remove one of the three accented syllables required to trigger the second. Thus Vanderslice and Ladefoged appear to predict that either the second or the third heavy syllable in such a phrase may be less prominent than the first, but not both. Again, this is surely not correct. Both *te* and *gra* in *seven telegraphic messages* may be less prominent than the initial syllable. Actually, I would go so far as to say they *must* be less prominent than the initial accented syllable in a well-formed phonological phrase. My theory predicts that the second and third strong syllables in the following phrases will both be perceived as less prominent than the first if the first is accented. This prediction appears to be correct.

- 41 a. 'seven 'tele'graphic 'messages
 b. 'Daddy 'tele'graphed as 'well
 c. 'unre'lia'bility 'hurts
 d. he 'plays the 'clari'net a 'lot

My theory also characterizes the second and third syllables in these examples as equally prominent. There is no distinction between accented and unaccented strong syllables between the first and last accented syllables in a phonological phrase, i.e. between the prenuclear accent and the nuclear accent. Elsewhere in their paper, Vanderslice and Ladefoged (1972:827) argue that there is no distinction between accented and unaccented syllables in postnuclear position: "all postnuclear heavy syllables are unaccented by definition". They point out-- correctly, I think-- that empirical evidence for such a distinction "is substantially non-existent" (p. 828). The same can be said of the accented/unaccented distinction in prenuclear position. Vanderslice and Ladefoged preserve this distinction, but they certainly do not refer to any empirical evidence to support it. I have already discussed evidence which bears on this question. My theory, unlike theirs, correctly predicts that *partyline*₁ and *partyline*₂ are homophonous both in postnuclear position and in prenuclear position.

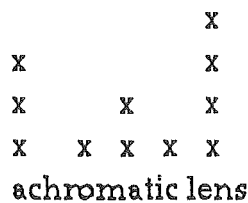
3.3. Liberman and Prince

Lieberman and Prince 1977 is the most recent treatment of variable stress to be discussed in this chapter, but it is not very recent. There is now a substantial literature on metrical phonology, building on the work of Liberman and Prince, which constraints of space and time force me to ignore here. Like Vanderslice and Ladefoged, Liberman and Prince propose to account for stress shift in some words by means of a rule designed especially for the purpose, i.e. an ad hoc rule. One of their examples is the phrase *achromatic lens*, in citation form, which I would transcribe as in 42.

42. 'achro'matic 'lens

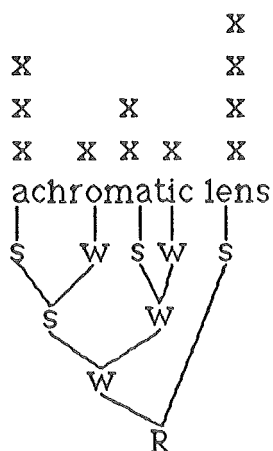
Using what they call a metrical grid, Liberman and Prince represent this phrase as follows.

42a.



In this sort of structure, the "degree of stress" of a given syllable is represented by the height of the column of marks that stands over it. Such a metrical grid is derived in a rule-governed way from a tree-like structure in which relative prominence is represented in a less perspicuous manner. The tree corresponding to 42a is shown in 42b.

42b.



Given this, Liberman and Prince and I agree that 42a is the right metrical grid for a citation utterance of *achromatic lens*. Unfortunately-- as they make clear-- without a Rhythm Rule, their theory would predict a different pattern of prominence for this phrase, viz. 43a.

43a.

```

                x
              x  x
            x    x
          x  x  x  x
    achromatic lens
  
```

Notice that in 43a, the third syllable of *achromatic* is more prominent than the first, as it is when this word is spoken by itself. This is because Liberman and Prince's accent rules, which construct the trees, preserve the relative prominence of the syllables in a word when that word is embedded in a phrase. They take the generalization that relative prominence is preserved under embedding to be fundamental, or linguistically significant. This does not mean that they believe it is always preserved, but that they treat as exceptional those phrases in which it is not.

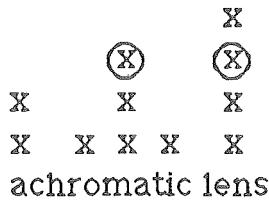
Liberman and Prince propose a Rhythm Rule, called Iambic Reversal, which converts representations like 43a into representations like 42a. Of course, it must be general enough to work for phrases like *thirteen men* and *Tennessee air*, which have slightly different patterns of prominence, but not so wide-ranging as to modify representations which are already "correct". Their rule is triggered by a "clash" of stresses, which is defined in terms of the elements which constitute a metrical grid (p. 314).

Elements are metrically *adjacent* if they are on the same level and no other elements of that level intervene between them; adjacent elements are metrically *alternating* if, in the next lower level, the elements corresponding to them are not adjacent; adjacent elements are *clashing* if their counterparts one level down are adjacent.

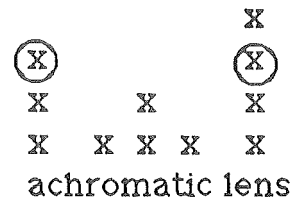
Given these definitions, the circled elements are metrically adjacent in both 43a and 42a. Those in 42a are alternating, since their counterparts one level down are not adjacent. In 43a, however, the circled elements are clashing,

since those immediately below them are adjacent. Therefore 43a is liable to be changed by the Rhythm Rule, while 42a is not.

43a

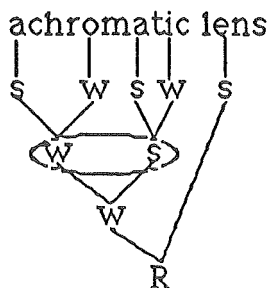


42a

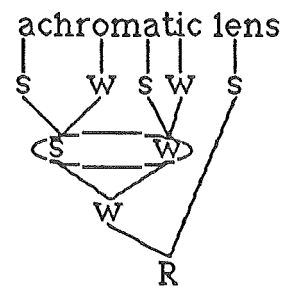


In this case, of course, the effect of the Rhythm Rule will be to change 43a into 42a. In fact, Liberman and Prince's rule operates not on metrical grids but on the relative prominence pattern as represented in a metrical hierarchy, or tree. That is, it changes 43a into 42a indirectly, by changing the tree structure in 43c as shown below. The metrical grid 42a is derived automatically from the modified tree in 42c.

43c



42c



Let us now consider some of the weaknesses of this approach.

First of all, Liberman and Prince (1977:320-1) make it clear that

Iambic Reversal is an optional rule, which nowhere mentions the notion "stress clash", but simply generates an optional relative prominence pattern in a wide variety of cases. In some of these, one option requires a scansion that contains a stress clash, while the other does not. . . . [T]he choice of a particular option in real-life situations depends on the independent consideration of such goals as maximization of metrical alternation, equalization or maximization of interstress intervals, etc.

Thus they claim that 42 is merely more likely than 43. The notion of optionality makes their hypothesis rather difficult to falsify. It is not at all clear how we are to test the claim that "one can always say" phrases such as 43.

Certainly, you can say anything you want, but this does not mean that all structures are well-formed. Compare this sort of claim with my theory's prediction that 43 is ill-formed. I would say that if the third syllable of *achromatic* is perceived to be more prominent than the first in *achromatic lens* and not more prominent than *lens*, then the utterance will be analyzed by the hearer as two phonological phrases. This has empirical consequences, especially when combined with a theory of intonation, and such a claim could be tested by means of replication experiments. We find support for this prediction in Liberman and Prince's (1977:320) observation that "the interstress interval is prone to be lengthened" in unretracted cases such as 43.

The claim that the rhythm rule is optional predicts not only that 43 may always occur, but also that 42 may always occur. This, too, is questionable, to say the least. As I pointed out with regard to Gimson, there are phrases which exhibit a stress clash as defined by Liberman and Prince, and in which the conditions for Iambic Reversal are met-- in particular, the syllable to be amended is not nuclear-- where the shift may not occur, such as *seven achromatic lenses*. Whether the shift is permitted here clearly depends on what precedes *achromatic*, as well as on what follows it. According to Liberman and Prince's formulation, what precedes *achromatic* has no effect on the operation of the rhythm rule, or rather, on the relative acceptability of the "retracted" form of the word. According to them, the crucial stress clash involves the syllable *mat* and a following syllable. No doubt they would claim that this was to be explained in terms of "maximization of metrical alternation" or some such principle. In this vague form, of course, such a claim is untestable, but this is not the point here. More significant is the fact that their theory would need to be complicated even further in order to achieve results which our much simpler theory produces naturally.

A more serious inadequacy of Liberman and Prince's theory is its prediction that-- apart from, say, utterances correcting mispronunciations-- either the first syllable or the third syllable of words like *achromatic* will be most prominent. If the rhythm rule applies, the first syllable will be more prominent than the third. If it does not apply, the third syllable will be more prominent than the first. Since Liberman and Prince's stress rules embody the

assumption that relative prominence is preserved under embedding, their theory does not predict that these two syllables can be equally prominent. We have already seen that they can indeed be equally prominent, and that the theory presented here correctly predicts the conditions under which this occurs. The third syllable will be the most prominent if, and only if, it is nuclear. The first syllable will be the most prominent if it is accented by the accentuation rules and the third syllable is not. When neither of the [+stress] syllables is [+accent], they will be equally prominent. This will happen whenever the word in question is either postnuclear or between the first accented syllable and the nuclear syllable.

The generalization which Liberman and Prince take as fundamental is, in my view, a spurious one. They illustrate the tendency for relative prominence to be preserved under embedding in this way (p. 251):

Thus the compound *whale-oil* (said in isolation) has its main stress on the word *whale*, with *oil* having some lesser degree of stress, and this inequality is felt to be preserved in the phrase *whale-oil lamp*, although the main stress of the phrase as a whole now falls on the word *lamp*.

In fact it is only because the first stressed syllable of *whale oil* is the primary one that the relative prominence is preserved. When an afterstressed word such as *Tennessee* or *cottage cheese* or *party line*₂ is embedded in the same way, the relative prominence of its syllables is not preserved. The fact that relative prominence is preserved for the great majority of English compound nouns is simply a consequence of the fact that the great majority of such words are forestressed. If the relative prominence is preserved for most English words, it would simply be a consequence of the fact that the first stressed syllable and the primary syllable were identical in most words.

4. Revisions

I have argued that the prominence features and accentuation rules outlined in section 2 offer a better account of variable stress than the alternatives discussed immediately above. While this theory is clearly up to the task in hand, it is far from perfect. I wish to conclude this chapter by

correcting some-- and only some, I stress-- of the more obvious inadequacies. In subsection 4.1 I revise the Nuclear Accentuation Rule, and consider some of the consequences of this change. In 4.2 I discard two redundant binary features, and make the minor rule changes necessitated by this move.

4.1. A new NAR

The Nuclear Accentuation Rule presented in section 2 predicts that only [+primary] syllables will be [+nucleus]. This prediction is correct if we confine our attention to phonological phrases in which only entire words are focussed. If we cast our net somewhat wider, however, the general statement represented by that rule is patently untrue, as the following examples show.

44. this whisky wasn't exported from Ireland | | it was **de**ported | |

45. it's not just anti**a**merican feelings | | it's anti**a**mericanism | |

46. I **di**dn't say it was possible | | I said it was **i**mpossible | |

Some linguists refer to this phenomenon as "repair"; a focus the speaker considers incorrect is being repaired. Others would describe it as metalinguistic, but it is clear that a theory which can account for it is preferable to one that cannot. We can account for it by revising the Nuclear Accentuation Rule so that it makes [+nucleus] the focussed syllable which has the most plusses.

Nuclear Accentuation Rule II

Make the most prominent [+focus] syllable of the last [+focus] word in a phonological phrase [+nucleus].

A "[+focus] word" is one containing material which is [+focus]. The "most prominent" syllable is the one which is positively specified for the greatest number of prominence features. Thus a syllable which is [+primary, +stress] is more prominent than a [-primary, +stress] syllable, which is in turn more prominent than a [-primary, -stress] syllable.

Let us see what this revised rule predicts about the accentuation of the eight-syllable word *antiamericanism*, given the following lexical prominence specification.

	anti a mericanism							
P	-	-	+	-	-	-	-	-
S	+	-	+	-	-	+	-	-

which are [+strength] whether they are accented or not. A much more natural solution is to suppose that those monosyllables which are strong only when accented are [-stress] in their underlying form. With the modified NAR, we are free to adopt this simple solution.

Permitting words with no stressed syllable will have consequences for the Prenuclear Accentuation Rule, because it refers to the feature [stress]. For ease of reference I repeat here the version proposed above.

Prenuclear Accentuation Rule II

- (i) If the first [+focus] word precedes the first content word, make the first [+stress] syllable of the first [+focus] word [+accent, +strength].
- (ii) If the first [+focus] word does not precede the first content word, make the first [+stress] syllable of the first content word [+accent, +strength].

The first part of the rule must be amended so that it can accent unstressed focussed syllables as well as stressed ones. This can be achieved by means of two ordered rules, only one of which may apply to a given phrase. This disjunctive pair of rules will accent the first stressed syllable of the first focussed word if it contains a stressed syllable, and the first syllable of that word otherwise. This will work as long as none of the stressless words has more than one syllable.

- (a) Make the first [+stress] syllable of the first [+focus] word [+accent].
- (b) Make the first syllable of the first [+focus] word [+accent].

These can be collapsed into a single rule with parentheses around [+stress]:

Prenuclear Accentuation Rule III

- (i) If the first [+focus] word precedes the first content word, make the first ([+stress]) syllable of the first [+focus] word [+accent].
- (ii) If the first [+focus] word does not precede the first content word, make the first [+stress] syllable of the first content word [+accent, +strength].

As far as I can determine, none of the stressless words are content words. This fact suggests that we might simplify the PAR by eliminating the references to content words. This is certainly an attractive possibility, but since further complications arise when one pursues it, I will not do so here.

In section 3.2 I pointed out that my two lexical prominence features were distributed differently from those of Vanderslice and Ladefoged (1972). We are

now in a position to see why my analysis is to be preferred. Briefly, their solution is inadequate because it predicts that [-stress] syllables, and post-primary [+stress] syllables, cannot be nuclear.

Recall that Vanderslice and Ladefoged's deaccentuation rule converts [+ACCENT] syllables into [-ACCENT] syllables if they are [+mentioned]. Their intonation rule then makes the rightmost remaining [+ACCENT] syllable [+Intonation]. Note that neither of these accent rules is capable of making [-ACCENT] syllables [+ACCENT]. From this it follows that these rules will not make syllables which are [-ACCENT] in the lexicon nuclear. Among such syllables would be the first of *impossible* and the penultimate of *antiamericanism*, which clearly can be nuclear. Of course, our NAR did not account for such phenomena until we revised it, but modifying it was a simple matter. Vanderslice and Ladefoged cannot adopt a version of NAR-II without abandoning the claim that, for example, the first and last syllables of *clarinet* are equally prominent in the word's underlying representation.

4.2. Redundant features

In another respect, Vanderslice and Ladefoged's theory is superior to the one I have outlined in the previous sections. Their theory is more economical in terms of features. They use three features, whereas as I have proposed five binary features of prominence-- three for describing phrasal prominence, and two for lexical prominence. This system is more economical than that proposed by Jassem and Gibbon (1980), which uses six features, but there is no doubt that I have been extravagant in postulating features. Only one set of three features is required. In mitigation, I would like to plead that the principal motivation for using two sets of features in this exposition was a desire for clarity.

The simplest solution is to eliminate the features [stress] and [primary], replacing them in the underlying form of syllables with the features [strength] and [accent], respectively. So instead of characterizing a given syllable as [+primary], we say that it is [+accent] at a particular level of description-- i.e. before the operation of the accentuation rules, or more generally, at the level of input to the phonological component. Using five features, we said that a

[+primary] syllable, such as the third one of *afternoon*, is not necessarily [+accent]. Using three features, we must say that a syllable which is [+accent] before the operation of the accentuation rules is not necessarily [+accent] thereafter. From the point of view of the revised feature system, I have heretofore used one term, e.g. [primary], when talking about the specification of a feature at one level of description, and another, e.g. [accent], when referring to its specification at a less abstract, more phonetic level. Again, using five features, we said that a [-primary] syllable, such as the first one of *afternoon*, is not necessarily [-accent]. Using three features, we say that a syllable which is [-accent] before the operation of the accentuation rules is not necessarily [-accent] after their operation.

This modification permits us to eliminate the Strength Rule, which applied following the accentuation rules in order to make [-accent, +stress] syllables [+strength]. On the other hand, it appears that a rule will be required to ensure that syllables which have not been accented by an accentuation rule are not [+accent] on the surface. If we let the accentuation rules as formulated apply to underlying forms described in terms of [strength] and [accent], they will make incorrect predictions about the number of accented syllables in many phrases. In particular, they will produce too many accented syllables. Instead of the one or two which we require, there will be at least one [+accent] syllable for each content word, because (a) we are supposing that each lexeme contains one [+accent] syllable, and (b) neither of our accentuation rules makes [+accent] syllables [-accent]. Clearly some sort of deaccentuation rule is needed.

If this new rule were to apply after all the accentuation rules, it would have to be global. That is, it would need the power to distinguish accented syllables created by accentuation rules from those which were not. Since this sort of power is undesirable, let us suppose instead that the following rule applies after the nuclear accentuation rule and before the prenuclear accentuation rule.

Deaccentuation Rule

Make all [-nucleus, +accent] syllables [-accent].

The Nuclear Accentuation Rule needs to know which syllables are [+accent] in the lexicon. The Prenuclear Accentuation Rule does not require this

information. It needs to know which syllables are [+strength] in the lexicon, and the Deaccentuation Rule does not affect the feature [strength].

An even more minor disadvantage to the elimination of redundant features is that it makes the term "accentuation rules" less appropriate. Now [+accent] syllables exist before these rules apply. Formerly no syllable was [+accent] except by virtue of the accentuation rules.

Summing up, we postulate the following accentuation rules, all of them obligatory, applying in the order given, prior to the tone association rules.

Nuclear Accentuation Rule (II)

Make the most prominent [+focus] syllable of the last [+focus] word in a phonological phrase [+nucleus, +accent, +strength].

Deaccentuation Rule

Make all [-nucleus, +accent] syllables [-accent].

Prenuclear Accentuation Rule (III)

- (i) If the first [+focus] word precedes the first content word, make the first ([+strength]) syllable of the first [+focus] word [+accent, +strength].
- (ii) If the first [+focus] word does not precede the first content word, make the first [+strength] syllable of the first content word [+accent].

This set of rules applies to representations of phonological phrases in which each syllable is specified for the binary prominence features [strength] and [accent]. No word contains more than one [+accent] syllable, and each [+accent] syllable is also [+strength].

I have shown that a theory incorporating these rules and features will account for the phenomenon of variable stress in a simple and natural way. Each of the rules applies to every phonological phrase, whether it exhibits a stress shift or not. This explanation contrasts with the three alternatives discussed in 3 above. All three fail to predict that the crucial syllables are equally prominent in some positions. More significantly, all three account for

variable stress by adding a rule to the grammar, a rule which accounts only for those phrases in which the most prominent syllable in a word is to the left of the syllable which is most prominent in the citation form of the word. This sort of approach is not only unnecessarily complicated, but also ad hoc.

Chapter 3

Obligatory Prenuclear Tones

In the tunes we have discussed so far, there is only one nucleus tone, and it is the first of the obligatory tones. That is to say, there are no obligatory tones before the nucleus one. This raises the question to which this chapter is devoted. Are there any tunes with an obligatory tone preceding the nucleus tone?

Others have answered this question in the affirmative, and I agree; but I think they have chosen the wrong tunes. O'Connor and Arnold (1973) analyze two tonal morphemes, or tone groups as they call them, as containing obligatory heads or preheads. I will argue that neither of these morphemes, the Low Bounce and the Long Jump, contains an obligatory tone before the nucleus tone. I will discuss the first of these in section 1, and the second, which appears to be the same as the surprise/redundancy contour of Liberman (1978:98-107), more briefly in section 2. In section 3, I will show that the tune commonly called a fall-rise, O'Connor and Arnold's Switchback, does contain such a tone. Unlike the obligatory prenuclear tones proposed by O'Connor and Arnold and by Liberman and Sag, this one is not an accent tone. The representation which I propose for the Switchback requires that the tone association rules be modified. In the final subsection I consider how this should be accomplished.

1. The Low Bounce

According to O'Connor and Arnold, the Low Bounce tone group has two forms (confining our attention to what they call unemphatic forms), as follows. (Parentheses indicate optional elements.)

- a. (Low Prehead +) High Head + Low Rise

b. High Prehead + Low Rise

Using curly brackets to indicate disjunction, we can represent this claim in our framework as follows.

$$\left\{ \begin{array}{c} (1)\underline{3} \\ 3 \end{array} \right\} \underline{13} \quad \text{Low Bounce (O\&A)}$$

This means that no instances of what they call the Low Rise nucleus --13-- with an initial nuclear syllable are tokens of this tune. All instances of their Low Bounce have at least one syllable before the nuclear syllable, and the syllable(s) immediately preceding the nuclear syllable must be high (i.e. attached to a 3-tone).

In their system, a Low Rise nucleus which is not preceded by any syllables is an instance of another tune, the Take Off. Their version of the unemphatic Take Off has the following structure.

(Low Prehead +) (Low Head +) Low Rise

As a tonal morpheme, it would look like this:

(1)(1)13 Take Off (O\&A)

The accent tone is actually redundant, but that need not distract us here.

If we simply want to classify phrases by tune in a consistent way, then O'Connor and Arnold's description of the headless Low Rise is as good as any. But if we want our description to be the right one, one which matches the native speaker's use of particular tunes, then their analysis is unacceptable. We shall see immediately below that low rises with neither head nor prehead are used in the same way as clear examples of the Low Bounce tune. If this is so we should revise the Low Bounce so as to accommodate such phonological phrases. This is easily done. I have already introduced the revised version in section 3.2 of Chapter 1, and I repeat it here.

(1)(3)13 Low Bounce (DH)

To find some facts which speak against O'Connor and Arnold's analysis we need look no further than their own examples. They give several examples of phrase-initial low rises where the nuclear syllable is an interrogative word such as *what*, *when*, *where*. "The questioner's tone is *wondering*," they say,

"as though he was mildly puzzled that such a question should have been asked or that he should have been given the information he was given"(p. 59).

1. The meeting's at five. ·When? | | (I thought it was six.)
2. How did he do it? ·How did he do it? | | (Perfectly obvious.)
3. His name was Scroggs. ·What? | | ·Scroggs? | |

They present these headless low rises, as I will call them, as examples of the Take Off. When discussing low rises preceded by higher syllables, that is, undisputed instances of the Low Bounce, they note that "when the nucleus is the interrogative word, the effect of repetition and puzzlement of the Take Off returns"(p. 64).

4. I saw him at Wembley. You 'saw him ·where? | |
5. They did it last week. They 'did it ·when? | |

In other words, the headless low rises have the same effect as the Low Bounces. It is important to note that the "puzzlement of the Take Off" was in fact exemplified not by clear instances of the Take Off tune but by the headless low rises quoted above.

Since O'Connor and Arnold organize their examples by sentence-type, we will turn now from questions to commands. "Commands with the Low Bounce," they say, "imply that the speaker is somehow . . . in a superior position to the listener, with the result that the speaker sounds encouraging and perhaps calmly patronising. For this reason these commands are frequently used to children"(p. 65).

6. 'Come to ·Daddy. | |
7. 'Blow your ·nose o·dear. | |
8. 'Don't ·worry. | |
9. 'Move a·long o·please. | |

This description also applies to all the examples O'Connor and Arnold give of putative Take Off commands in which the first syllable is nuclear (pp. 62, 144-5, 148). Indeed, they indicate explicitly that the majority of them are to be imagined spoken by parents to children or by teachers to students. The

following examples are typical (p. 148).

10. [Father to small son who is riding his bicycle a little too fast] ·Slowly.
11. [Mother to small son who is teasing a puppy] ·Gently. ·Careful.
12. [Starter to sprinter who has beaten the gun] ·Wait for it.

Since these appear to have the same pragmatics as clear examples of the Low Bounce, it is reasonable to analyze them as examples of that tonal morpheme rather than the Take Off. This analysis of headless low rise commands becomes even more attractive when we see that they are pragmatically quite different from commands in which a low head precedes the low rise, i.e. undisputed examples of the Take Off. When such a tune is used with a command, say O'Connor and Arnold (1973:61), the effect is of "appealing to the listener". Their examples include the following (p. 157):

13. ·Be a ·sport.
14. ·Don't ·worry about ·that.
15. Oh ·do hurry ·up.

Analyzing the headless low rises above as Low Bounces accounts for the fact that they function like Low Bounces and for the fact that they do not function like Take Offs. O'Connor and Arnold's description of all headless low rises as examples of the Take Off is at odds with these facts.

Our final group of examples is non-final phrases. O'Connor and Arnold say that "the Low Bounce is frequently used with non-final groups, when the speaker is leading up to something more. The effect is to create expectancy" (p. 63). This is also a good description of the use and effect of many phrases beginning with a low rise, like the following:

16. ·one || ·two || ·three || ·four || ·five ||

This is presented by O'Connor and Arnold to illustrate not the Low Bounce but the Take Off, "used for continuative purposes, to show that there is more to be said" (p. 58). It is a shame they could not illustrate this with a certain instance of the Take Off, that is, a low rise preceded by a low head. Had they used bigger

numbers, they would have discovered clear examples of the Low Bounce used for enumeration:

17. 'twenty -one | | 'twenty -two | | 'twenty -three | |

It is unreasonable to suppose that we use one tonal morpheme for *twenty three* and a quite different one for *three*. The obvious solution is to assume that the high accent tone of the Low Bounce is optional. This allows us to label all the headless low rises quoted above as Low Bounces, which will explain why they are used in the same way.

I have argued that some phrases which begin with a nuclear syllable and carry a low rise pitch pattern are examples of the Low Bounce. This is to disagree with O'Connor and Arnold, whose description of the Low Bounce excludes such phrases. I have not argued that all such phrases are examples of the Low Bounce. Some of them are, as O'Connor and Arnold claim, examples of the Take Off. Others which have been used to illustrate the Take Off are, as I shall argue in Chapter 8, examples of the so-called contradiction contour. According to my analysis, then, some realisations of the Low Bounce are homophonous with some realisations of the Take Off (and, possibly, with some realisations of the Contradiction Contour). That is, headless low rises are ambiguous.

O'Connor and Arnold's Low Bounce tune contains an obligatory 3-tone before the nuclear tone. We have seen that, insofar as their description of the Low Bounce amounts to a testable claim, it is incorrect. A revised Low Bounce which has an optional 3-tone before the nuclear tone accords better with facts about certain kinds of phrases. We conclude that the Low Bounce does not in fact refute the hypothesis that the nuclear tone of a tonal morpheme is always the first obligatory tone.

The evidence against O'Connor and Arnold's case is not obscure. On the contrary, there is plenty to be found among their own examples. Since they were familiar with the relevant facts, we are entitled to ask why their analysis is so obviously at odds with them. The answer is relatively simple. I pointed out above that, once we revise the Low Bounce, a 13 phrase could be either a Low Bounce or a Take Off. It would not be possible to decide, on the basis of phonetic information, which tonal morpheme is being used. This is not a

possibility which O'Connor and Arnold were prepared to consider. With their system, there is never any doubt, given the phonetic data, about which tone group is exemplified. The transcriber's job-- or the computer's-- is much simplified. This is comparable to an insistence on deciding on the basis of phonetic data alone what word [rid] represents. Is it the same as the word that appears in *reading*, or the same as the word that appears in *reedlike*? If the native speaker of English does not always know, we do not want a theory which is always certain. O'Connor and Arnold's theory of intonation is just such a theory, driven by the desire for correct transcription. It prohibits homophony. What we want, surely, is a theory that accounts for ambiguity, instead of concealing it.

It appears that O'Connor and Arnold's acceptance of the following condition, whose wording is due to Halle (1959:21), prevents them from considering what seems to us the obvious solution to the headless low rise:

A phonological description must include instructions for inferring the proper phonological representation of any speech event, without recourse to information not contained in the physical signal.

I will have more to say about this constraint in my discussion of putative binuclear tunes in Chapter 7.

2. The Long Jump

We will find further support for this explanation of O'Connor and Arnold's analysis if we turn to the second type of tone group purported to contain an obligatory tone before the nuclear tone. This is the Long Jump. Like the Low Bounce, it is identical with another tune from the nuclear tone to the end. Recall that their Low Bounce has the same nuclear tone-- Low Rise-- as the Take Off. The Long Jump has the same nuclear tone-- High Fall-- as the High Drop. The prenuclear portions differ. Where the High Drop has an optional High Head, the Long Jump has a Rising Head, which is obligatory.

(1)(3)31 High Drop (O&A)

(1)1 31 Long Jump (O&A)

In this representation of the Long Jump, the Rising Head is replaced by an

accented 1-tone. This simplification does not affect the argument. If both of these tone groups had optional accent tones, then some examples of the High Drop would be homophonous with some examples of the Long Jump. In particular, it would not be possible to determine the correct labelling of a phrase beginning with a High Fall nuclear tone, i.e. 31 in my terms. O'Connor and Arnold obviate such a possibility by making the prenuclear tone obligatory in the Long Jump. Thus any phrase beginning with a High Fall nuclear tone will be a High Drop, by fiat. The obligatory prenuclear tone in the Long Jump, like the one in the Low Bounce, is simply an artifact of their horror of homophony. If we survey the tone groups in their book, we find that the two groups which contain obligatory prenuclear elements both have the same nuclear tone as another tone group. And these are the only pairs of tone groups which share nuclear tones.

Now that we have seen why O'Connor and Arnold's Long Jump has an obligatory prenuclear tone, it would be superfluous, and straining the reader's patience, to demonstrate at length that a phrase beginning with a nuclear syllable can be used in the same way as undisputed instances of the Long Jump. It seems unlikely, to say the least, that deciding to utter an emphatic *why?* instead of *what for?* would require a speaker to select an entirely different tonal morpheme. The hypothesis that the choice of tonal morpheme is independent of the number of syllable in a phrase appears to be a fundamental one in studies of intonation. In the absence of reasons to abandon this assumption, I will suppose that the Long Jump does not contain an obligatory prenuclear tone.

3. The Switchback

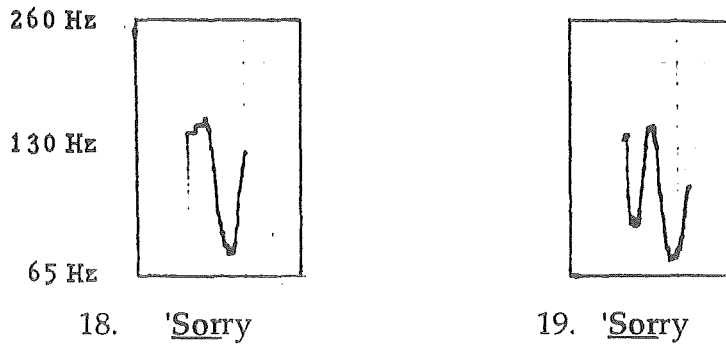
We have now dealt with two potential groups of counterexamples to the hypothesis that the nuclear tone is the first obligatory tone of a tonal morpheme, and have dismissed both of them. There is another English tonal morpheme which has not been proposed as a counterexample to our hypothesis, but is. This is the tune which O'Connor and Arnold call the Switchback. In Halliday's treatment it is Tone 4 with a high pretonic. I will

argue that the best representation for this morpheme is (1)(3)1313. In section 3.1 I will compare this morpheme with the Dipper morpheme, since the two are sometimes confused. In section 3.2 I will consider some alternative hypotheses about the form of the Switchback, and briefly indicate their inadequacies. Since the application of the tone association rules to this morpheme leads to complications, their operation will not be discussed until section 3.3.

3.1. Switchback and Dipper

The Switchback morpheme, which I will argue has the form (1)(3)1313, is similar to the Dipper, which I analyzed in Chapter 1 as (1)(3)313. The last three tones are identical, and so are the optional tones, in my analysis. In view of this it is perhaps unsurprising that not all treatments of intonation distinguish the two. The term Fall-Rise is appropriate for the Dipper, and Rise-Fall-Rise for the Switchback, but many discussions of intonation refer only to fall-rises, in spite of the fact that the Dipper, in my experience at least, is more marked and less common than the Switchback. I will not attempt to document the latter claim, but it is supported by O'Connor and Arnold's inclusion of the Switchback but not the Dipper in their textbook. I do not doubt that there are dialect differences, but neither do I doubt that many references to fall-rises in the linguistics literature are references to the Switchback.

Halliday (1967) is one of the linguists who do keep the two morphemes apart. The Switchback can be identified with Halliday's Tone 4, while the Dipper corresponds to a variety of his Tone 2, as I pointed out in Chapter 1. In my dialect, too, there are clear formal and functional differences between the two tunes. Consider the following minimal pair. In each phrase the first syllable is nuclear.



In the Switchback (19), but not the Dipper (18), the nuclear syllable must begin low. The pitch rises before it falls, even when the first sound is voiceless, as it is here. We account for this by postulating an obligatory 1-tone before the nucleus 3-tone. The difference between the tunes also involves the syllable immediately before the nuclear syllable. It, too, is low, unless it is accented, in which case it is falling, i.e. high at the beginning and low at the end. I will illustrate these effects a little further on. I will argue that this characteristic can also be attributed to the pre-nuclear 1-tone, so that the presence of this tone is the only thing which distinguishes the underlying form of the Switchback from that of the Dipper.

There are functional differences corresponding to the formal differences. The pragmatics of the two tunes are distinct. For example, *Sorry* with the Dipper is an apology. With the Switchback, however, it is appropriate not as an apology, but as an echo, repeating an interlocutor's *sorry*. One can imagine two kinds of situations in which a speaker might echo an apology. In the first, she cannot understand why the other person is apologizing: "*Sorry.*" *What are you saying that for?* In the second kind of situation, she perceives the misdeed to be too serious to be repaired by a mere "sorry". Boosting, or increased pitch range, would be more appropriate in the latter scenario than in the former.

O'Connor and Arnold characterize the Switchback as a combination of a Falling Head and a Fall-Rise Nucleus. They do not discuss the Dipper, but it could certainly be characterized as a combination of a Fall-Rise nucleus with a High Head, which occurs in other tone groups. O'Connor and Arnold do not rule this out, carefully qualifying their statement that "the Fall-Rise always has the falling head before it" with the phrase "in this book". I suspect that they deliberately excluded this tune from their textbook to avoid unnecessary complications for the students, the teachers, and themselves. Had they

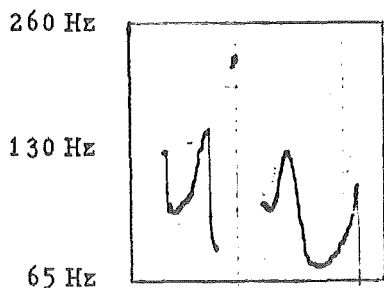
included both tunes, they would have felt obliged, for reasons discussed above, to designate one of the heads as obligatory. This would be a difficult decision because, as we have just seen, both have optional heads-- they contrast even when the nuclear syllable is phrase-initial. By omitting one tune, O'Connor and Arnold sidestep this problem. Deciding which tune to leave out was not so difficult, because the Switchback is the more common of the two.

3.2. Competing hypotheses

Let us now consider some of the representations which have been or might be proposed for the Switchback, with a view to pointing out their flaws.

3.2.1. It is fitting to begin with the version of O'Connor and Arnold, which converts into our framework as (roughly) (1)(321)313. There are several problems with their representation which I wish to discuss.

First of all, the term Falling Head is deceptive in that, as O'Connor and Arnold (1973:21) observe, "If there is only one syllable in the head, that syllable is high and level". That is, the pre-nuclear accented syllable is attached only to a 3-tone. Phonologically, this appears to be correct. Phonetically, of course, the pitch often rises to this level, as in the following display, but this is neither contrastive nor peculiar to this tune. Certainly the pitch of the first accented syllable is not falling.

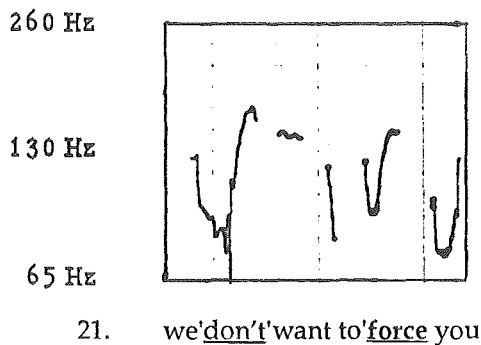


20. I'm 'not 'cold

A second problem concerns the character of the fall. O'Connor and Arnold's description of the difference between the Falling Head and the ordinary High Head (represented in our system as a parenthesized accent 3-tone) is not accurate. According to them, the first syllable of the Falling Head

"is rather high in pitch and any following syllables gradually carry the pitch lower"(p. 20). In the High Head, they say, "all the syllables are said on the same rather high pitch"(p.19). This is correct phonologically, but not phonetically, because the pitch ordinarily starts high and gradually descends, as it does in the Falling Head. This phenomenon is familiar to students of African tone languages as *downdrift* (see Hyman 1975:226). It is also familiar to most students of English intonation. For example, Crystal's recognition of downdrift is reflected in his decision not to mark for pitch-range a syllable which is slightly lower than the preceding syllable (1969:144).

We are talking here of Switchback phrases which contain a high prenuclear accented syllable, and more than one syllable between it and the nuclear syllable. There may well be such phrases in which the pitch descends on a steeper slope than a downdrifting High Head, but gradually. In my dialect, however, the descent is typically gradual up to about the beginning of the last prenuclear syllable, and then sudden. The syllable immediately preceding the nuclear syllable in the following example is best described as low.



We can account for this by supposing that the prenuclear syllable is attached to a 1-tone, and the three preceding it to a 3-tone. If the pitch descends gradually before that syllable, we can attribute this to downdrift. The Falling Head looks like an unnecessary complication.

This brings us to yet another problem with a representation of the Switchback in which the nucleus 3-tone is the first obligatory tone. It predicts that there will be no difference between the Dipper and the Switchback in phrases beginning with a nuclear syllable. This prediction is incorrect. Examples 18 and 19 above show that these two morphemes are phonetically

distinct in such phrases. When such phrases carry the Switchback tune, the first syllable starts low. We account for this by postulating that the 1-tone before the nucleus tone is obligatory.

3.2.2. We have now argued for a representation something like (1)(3)1313.

This is by no means the only possibility which avoids the flaws of O'Connor and Arnold's hypothesis. We might, for instance, suppose that it is not the 3-tone which is the nucleus tone, but the 1-tone, so that the obligatory tones are 1313. This appears to be how Leben (1976:81) would analyze this tune. In his system it would be L* H L, with the initial L starred for attachment to a nuclear syllable, plus "comma intonation", which "adds a H to the last syllable of a clause". The most obvious advantage of such a solution is that it preserves the hypothesis that the nucleus tone is the first obligatory tone in every tonal morpheme. The most obvious disadvantage is that, unless we make fundamental changes to the tone association rules, it makes incorrect predictions about the distribution of the tones of this morpheme. Consider, for example, a phrase with three postnuclear syllables, such as *Frank is coming*. The representation under discussion seems to predict that the nuclear syllable will be low and the following syllable high, by the following derivation.

'Frank is 'coming

1 3 1 3

ITA Rules

'Frank is 'coming

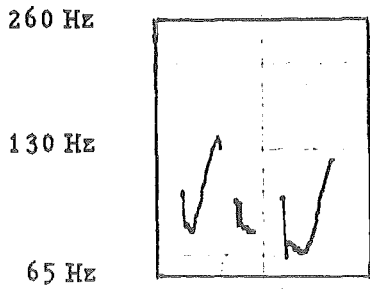
1 3 1 3

WF Rules

'Frank is 'coming

1 3 1 3

Both of these predictions are incorrect. The following display shows that the nuclear syllable is low at the beginning, but high at the end. The syllable immediately after it is low, not high.



22. 'Frank is 'coming

This corresponds to a representation like 22a in which the nuclear syllable is attached to both a 1-tone and a 3-tone, in this order, and the first postnuclear syllable to a 1-tone.

22a. 'Frank is 'coming
 1 3 1 3

3.2.3. We could achieve this result by making both the first and the second obligatory tones nuclear: 1313. Such a representation would save our hypothesis that the first obligatory tone is nuclear. It would survive, though, at the expense of the hypothesis that each tonal morpheme contains only one nucleus tone. The latter hypothesis is more fundamental, I think, than the former, but its loss is not the only disadvantage to this solution.

The second problem also applies to Leben's hypothesis discussed immediately above. If we make the 1-tone nuclear, it is difficult to use this tone to account for the drop in pitch before a non-initial nuclear syllable. Suppose that the following is the output we want for *we don't want to force you*.

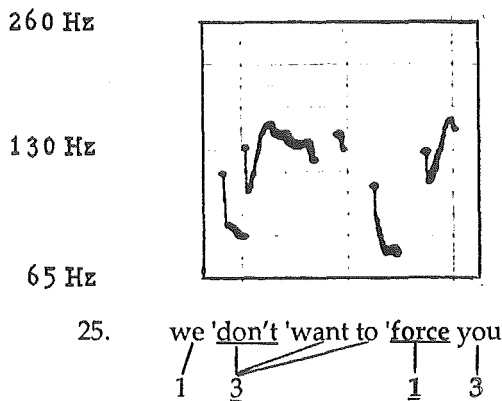
23. we 'don't 'want to 'force you
 1 3 1 3 1 3

We might try to achieve this by means of an assimilation, or flop rule, like the one referred to in section 1.3.3 of Chapter 1. This rule would come into operation after the 1-tone has been attached to the nuclear syllable, and would link it to the syllable to the left of the nuclear syllable. There are at least two difficulties raised by such a rule. In order for it to produce 23, the flop rule would have to apply before the WF rules, which will attach all the free interaccentual syllables to the first 3-tone. If the WF rules apply before the flop rule, the result will be 24, which incorrectly predicts that the prenuclear syllable

will be high at the beginning.

24. we 'don't 'want to 'force you
 1 3 1 3 1 3

The flop rule solution combined with a nucleus 1-tone runs into at least two difficulties. The first has to do with applying the flop rule before the WF rules. This flies in the face of the hypothesis, which appears to be well-founded, that the WF rules apply before flop rules. The second difficulty is that it makes incorrect predictions about the pitch contour of Low Bounce phrases. Recall that the Low Bounce contains, among other things, an optional accent 3-tone followed by a nucleus 1-tone: (1)(3)13. So does the Switchback according to the hypothesis under consideration here: (1)(3)1313. The flop rule cannot be allowed to apply to the Low Bounce, because there is never a low syllable between the accented syllables in a Low Bounce phrase. The pre-nuclear syllable in this position is always high, i.e. attached only to a 3-tone, as in 25.



Since the environment is phonologically the same, one could only prevent the flop rule from applying in the most ad hoc way.

3.2.4. It might be suggested that we do not need to use the 1-tone to account for the pre-nuclear drop. This could be done by adding an optional 1-tone in front of the nuclear 1-tone, giving something like (1)(31)1313. Obviously it would be uneconomical to use two separate tones to account for one stretch of low pitch, but adopting this kind of separation would not be merely uneconomical. To adopt this kind of representation would be a serious mistake, because it

amounts to a claim that the fall in pitch before the nuclear syllable and the rise on that syllable are unrelated. Using two separate tones implies that the identity of the tones is accidental, that the following hypothetical morpheme, in which the prenuclear optional tone is high rather than low, is just as likely or natural as the Switchback: (1)(3)1313. I find this difficult to believe. I find it easier to believe that the drop and the rise at the nuclear syllable are related. For this reason I wish to attribute them to the same source-- a single 1-tone.

3.3. Tone association rules

We are left with (1)(3)1313 as the best representation of the Switchback morpheme. But obstacles remain in our path. In particular, the set of tone association rules hypothesized in Chapter 1 will need to be revised. That this should be necessary is not surprising, given the state of our understanding of English tone association rules. Neither does this reflect badly on this particular representation. The other possibilities discussed above would also require us to amend the rules. Moreover, the changes required are not major.

If the tone association rules we have been supposing are applied to our representation of the Switchback and a phonological phrase such as *I'm 'not e'xactly 'cold*, the prenuclear 1-tone will be linked by the well-formedness rules to all of the interaccentual syllables, thus:

I'm 'not e'xactly 'cold

1 3 1 3 1 3

ITA Rules

I'm 'not e'xactly 'cold

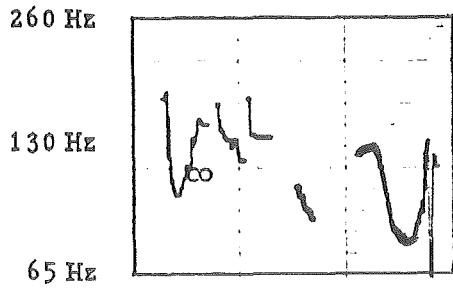
1 3 1 3 1 3

WF Rules

I'm 'not e'xactly 'cold

1 3 1 3 1 3

This is clearly not what we want. Only the last of the prenuclear syllables is low, as the following display shows.



26. I'm 'not e'xactly 'cold

Let us suppose that the correct representation of 26 is as follows.

26a. I'm 'not e'xactly 'cold
 1 3 1 3 1 3

The prenuclear syllable is attached only to a 1-tone. The syllables preceding it are attached to a 3-tone. The obligatory 1-tone is not attached in this representation to the nuclear syllable. For the sake of simplicity let us assume that this is the output we require from the tone association rules.

We will get this result if we suppose that the tone in question is linked to the prenuclear syllable before the WF rules come into effect. If this is so, then it must be attached by an initial tone association rule, since we are assuming that only these rules apply before the WF rules. Recall that we are supposing that two types of ITA rule are used in a grammar of English. Nucleus tone and accent tones are attached by rules of the first type, which attach marked tones to syllables marked in the same way. End tones are attached by a rule of the second type. It attaches a tone to the rightmost syllable in the phrase.

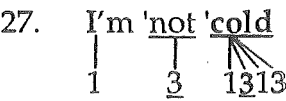
The rule which attaches the prenuclear obligatory tone of the Switchback fits neatly into the second category of ITA rules. This tone is attached not to the rightmost syllable of the phrase but to the rightmost prenuclear syllable. This rule, which we can call English ITA Rule 2b, will have to find the nuclear syllable and treat it just as English ITA Rule 2a treats the end of the phrase. If we assume that this rule applies to any obligatory prenuclear tone, it will not be necessary to mark the tone with a diacritic feature.

English ITA Rule 2b

Associate an obligatory prenuclear tone with the rightmost prenuclear syllable.

There is another difference between this rule and ITA Rule 2a, which attaches end tones. An end tone gets linked to the final syllable whether that

syllable is accented or not, i.e. whether it is attached to a nucleus tone or not. The prenuclear 1-tone of the Switchback, on the other hand, is not attached to the rightmost prenuclear syllable if that syllable is accented, as in *I'm 'not 'cold*. We saw in section 3.2.1 above that this phrase should be represented as in 27.

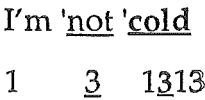


We can account for this by (i) supposing that ITA Rules 1a and 1b apply before ITA Rules 2a and 2b, and (ii) adding to Rule 2b the condition "if the syllable is free", i.e. not associated with a tone.

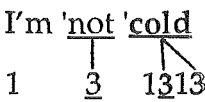
English ITA Rule 2b (Revised)

Associate an obligatory prenuclear tone with the rightmost prenuclear syllable, if it is free.

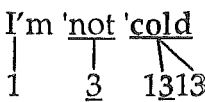
The derivation for 27 will therefore proceed as below.



ITA Rules



WF Rules 1-3



At this point in the derivation, the tone which concerns us is stranded, or floating, as is the penultimate tone. They will not be attached by the ITA rules or by the three universal WF rules. In section 3.3 of Chapter 1, I proposed that English has a fourth WF rule, which I repeat here.

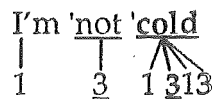
WF Rule 4

If after the operation of the first three WF rules there remains an unassociated tone and no unassociated syllables, associate the tone with the syllable attached to the tone on its right, if there is one. If there is no tone to the right, associate the tone with the syllable attached to the tone on its left.

This rule will attach both of the 1-tones left floating above to the nuclear

syllable. In both cases, the nuclear syllable is the syllable which is attached to the tone on the right of the stranded tone. The derivation therefore ends in this way:

WF Rule 4



The prenuclear 1-tone will also be stranded and then correctly attached by WF Rule 4 when the nuclear syllable is initial, as in the following derivation.

ITA Rules



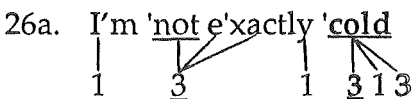
WF Rules 1-3



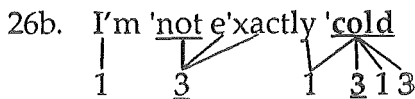
WF Rule 4



Let us return briefly to phrases with syllables between the first accented syllable and the nuclear syllable. In a phrase like this, the ITA Rules will attach the prenuclear 1-tone only to the prenuclear syllable.



This predicts that the nuclear syllable need not be low at the beginning, and this is certainly true, as the trace for 26 above shows. The nuclear syllable is not always phonetically high at the beginning in such a phrase. It may therefore turn out that 26b is an appropriate representation for some Switchback utterances of *I'm 'not e'xactly 'cold*.



It seems best to handle this kind of variation by means of an optional flop rule which applies after the WF rules. Such a rule will make a copy of the association line between the prenuclear syllable and the 1-tone in 26a, and let it

"flop" to the right, so that the 1-tone is associated with the nuclear syllable as well, as in 26b.

I have argued in this chapter that neither the Low Bounce nor the Take Off tonal morphemes contain an obligatory tone before the nucleus tone, but that the Switchback does. In order to account correctly for the distribution of its tones, it was necessary to postulate a fourth English ITA rule, namely 2b. This solution is certainly not problem-free, and further investigation of the alternatives is certainly warranted. It may turn out, for instance, that the double nucleus tone analysis of section 3.2.3 is preferable. As I pointed out, however, a very strong case would have to be made for it, because it is contrary to the basic and productive hypothesis that no tonal morpheme contains more than one nucleus tone. Since the field of tone association rules in English is largely uncharted territory, it is not surprising that we should encounter some problems. The mere existence of complications does not seriously threaten the claim that the Switchback contains an obligatory prenuclear tone. Our understanding of tone association rules is too hazy to provide solid support for arguments against particular representations if they are otherwise well motivated. It is clear, finally, that deeper investigation of tonal morphemes requires us to make our hypotheses about the phonological rules more precise than they have been heretofore, and this cannot be bad.

Chapter 4

The Fifth Tone

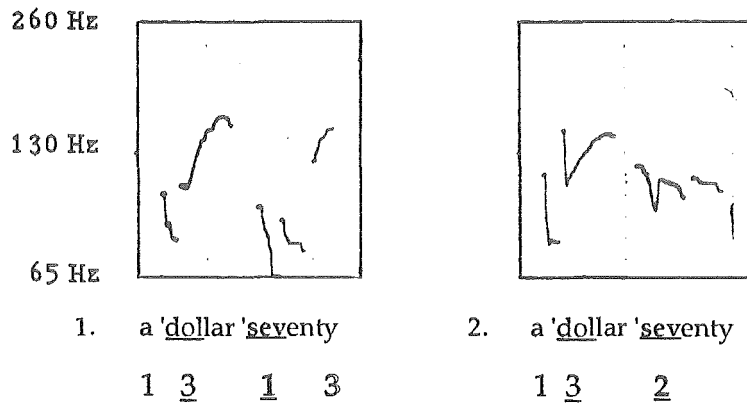
Most of the linguists who analyze English tunes into pitch levels agree that the number of levels required is four. Trager and Smith (1951), Pike (1945), and Wells (1945), for example, all postulate four pitch phonemes. A four-tone theory is attractive because a set of four tones can be analyzed with only two binary features. In this work I use five tones, or pitches. This chapter is devoted to the demonstration that four tones do not suffice. I will begin by reviewing the facts which lead me to postulate the fifth tone. Since so many linguists agree that only four tones are required, I will then look at the above-mentioned works with a view to explaining how they manage with fewer tones than we need.

1. Need for five tones

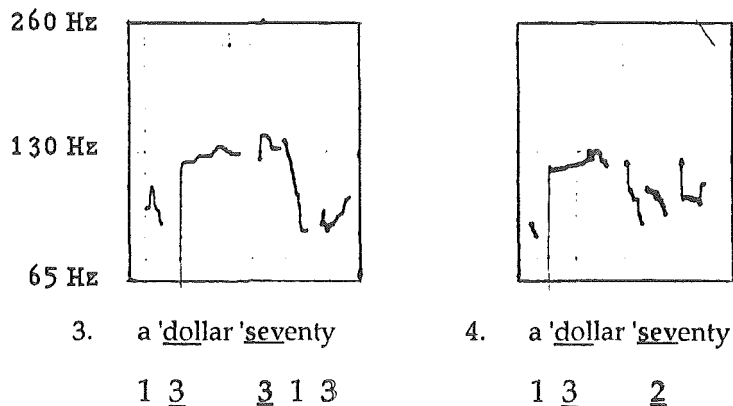
It is a simple matter to demonstrate that five tones are required. First I will show that we need two tones lower than the common prenuclear accent tone, which must therefore be tone 3. Then I will show that we also need two tones higher than this tone. Since there are no derivations in this chapter, the tone association lines can be omitted from the diagrams herein. No confusion should result if each tone number is placed under the first syllable associated with it.

O'Connor and Arnold discuss a tune which they call the Terrace. It consists of an optional low prehead, an optional high head, and an obligatory mid-level nuclear tone. I represent this tonal morpheme as (1)(3)2. The nucleus tone is lower than the accent tone but higher than a 1-tone. This is

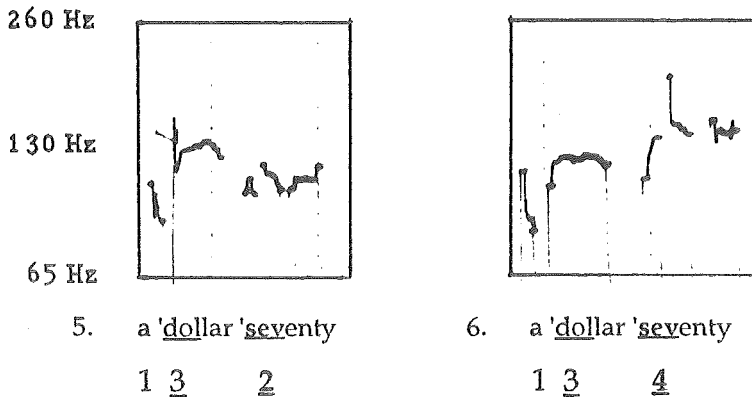
clear if we compare this tune-- on the right below-- with the Low Bounce on the left.



The nuclear syllable follows the second gap in the traces. A comparison between Terrace and Dipper also shows that the final tone of the Terrace-- on the right below-- is lower than the accent tone and higher than the penultimate 1-tone of the Dipper.



Tone 4 is required for the High Level tune. O'Connor and Arnold do not discuss this one, but Cruttenden, among others, does. I represent this tune as (1)(3)4 because the nuclear syllable is always higher than a preceding accented syllable, which, as we have seen, is attached to a 3-tone. Below I illustrate the High Level-- on the right-- compared with the Mid Level, or Terrace. Clearly these tunes are the same up to the nucleus tone. Because these examples were spoken fairly quickly, and the nuclear syllable is short, the pitch associated with the 4-tone is not reached until the end of the nuclear syllable. This is presumably a relatively superficial, i.e. phonetic, phenomenon.

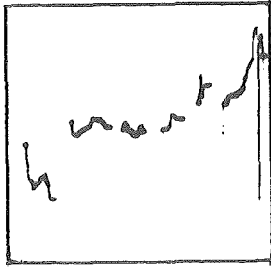
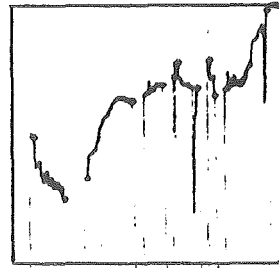


Tone 5 is required for the High Rise, where it is the final tone: (1)(3)45. It is marked as an end tone, which indicates that it is attached by English ITA Rule 2a to the last syllable of the phrase. I represent the nucleus tone of this morpheme as a 4-tone. That is, I claim that the nuclear syllable of a phrase associated with this morpheme must be higher, phonologically, than the preceding accented syllable. Thus it differs from the nuclear syllable in a Fall or Fall-Rise phrase, which may be higher but is not necessarily so. I represent the nucleus tone in those tunes as a 3-tone. Making the nucleus tone of the High Rise a 4-tone helps us to explain why people often run out of room when producing a High Rise phrase. In many utterances which are understood and intended as High Rises, the final syllable is not noticeably higher than the penultimate one, because the speaker has reached the top of her pitch range. Representing the nucleus tone as a 3-tone would make it difficult to account for this phenomenon, since the 3-tone is in use virtually all the time. If the nuclear syllable must be higher than this, it is not so surprising that speakers sometimes go too high too soon. My representation also predicts that if the first accented syllable of a phrase is nuclear, like the third syllable of *Are you coming to the party tonight*, it will be attached to a 4-tone. That is, it will be higher, *ceteris paribus*, than the third syllable of *Are you coming to the party tonight*, which is non-nuclear and attached to a 3-tone. This is in fact what happens, as we see below. (The third gap is the beginning of *party*. As in example 6 above, time is required to reach the 4-tone pitch.)

260 Hz

130 Hz

65 Hz

7. are you coming to the party to 'night1 3 4 58. are you coming to the party to 'night1 4 5

My claim that the High Rise contains a 5-tone does not, however, depend on the correctness of my representation of its nucleus tone. If we compare a High Rise (9) with a High Level (10) in the same register, it is clear that the final syllable of the former is considerably higher than the final syllable of the latter, which is attached to a 4-tone.

260 Hz

130 Hz

65 Hz

9. a dollar seventy1 3 4 510. a dollar seventy1 3 4

2. Trager and Smith

If four tones are not sufficient, we might well ask why so many linguists agree that they are. For Trager and Smith (1951) the answer to this question is rather complicated. In order to make it more comprehensible, I will first compare their approach with mine, explaining how and why my theory differs from theirs in various respects.

2.1. Comparison

In the first place, it must be said that there are substantial differences. Many of these are due to differing views on the aims of linguistics, which are

due in turn largely to differing views of the nature of scientific investigation. We agree that linguistics is a science, but Trager and Smith's science is positivist, while mine is Popperian. Their goal is classification; mine is explanation. For the most part, I will ignore differences which seem to be a consequence of these larger methodological differences. I will concentrate on similarities and on those differences which would remain if Trager and Smith's notions were incorporated into a theory of English which was intended as a model of linguistic knowledge (in the sense of Smith and Wilson 1979: 32-38).

There is very little difference between our tones and the pitches of Trager and Smith. Both are phonological units of a kind different from vowels and consonants. Some linguists seem to have been confused by their use of the word *phoneme* for these pitch-levels, but as Trager (1964) makes clear, they call all phonological units phonemes. As in our theory, the actual pitch of a tone/pitch varies. In particular Trager and Smith recognize four allophones of pitch height for each of their pitch phonemes. There are four allophones because they distinguish four stress phonemes, and claim that pitch height varies directly with stress. We do not adopt this, because it is incorrect. This is especially obvious in phrases where the nuclear syllable is low in pitch. It is more likely to be lower than non-nuclear syllables attached to the same tone than higher, as Trager and Smith predict. This is not to say that prominence is irrelevant to the height of a tone, but that the relationship is not so simple as they claim. It is also clear that there are factors other than stress or accent which determine the pitch of a tone.

Unlike Trager and Smith, we postulate general rules for the association of tones and syllables. Trager and Smith have no explicit rules in their grammar, but they certainly recognize that, in their own words, "Intonation patterns show allomorphs involving the scope of each of the pitch phonemes-- that is, the extent of the material included under each pitch" (1951:60). A complete Trager and Smith grammar would list each of these allomorphs along with the phonological context(s) in which it occurs. If we interpret a grammar as a model of a person's linguistic knowledge, this is a claim that the native speaker learns a large number of suprasegmental allomorphs just as she learns the

words of her language. This is unlikely, to say the least. There are generalizations to be made about the scope of tones/pitches, and without some kind of tone association rules, Trager and Smith cannot make them. We can account for the variations in scope by means of a few simple and general tone association rules operating on a small number of morphemes and an infinity of phonological phrases.

We agree with Trager and Smith that "intonation patterns depend for their scope and exact form on the phrase-superfixes" (1951:57), the latter being sequences of stress phonemes. That is to say, tone association rules (which account for the distribution of the tones of a morpheme) refer to the prominence features of syllables in the phonological phrase. Like Trager and Smith, I distinguish four degrees of prominence. They do this by means of four stress phonemes, while we use three binary features. This similarity is deceptive, though, because their four degrees of prominence do not correspond exactly to mine. In particular, the secondary stress of Trager and Smith is not the same as my non-nuclear accent. In the first place, they describe the stress on the first syllable of *psychological* (1951:50) or *operation* (1951:39) as tertiary, not secondary. Secondly, they permit secondary stress phonemes after the primary stress in a phrase. Our theory predicts that this will not happen, that there is no contrast between accented and unaccented after the nuclear syllable. Their examples of contrasts between secondary and tertiary stress in this position are unconvincing:

bláck + bîrd

bláck + bôard

Whíte + Hòuse (he lives in a) whíte + hône (not a brown one)

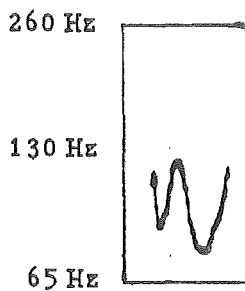
Their single example of tertiary in contrast with secondary stress before a primary stress is also unconvincing:

òld + máid "spinster"

ôld + máid "former servant"

Our means of representing what Trager and Smith call the scope of a pitch phoneme allows one tone to be associated with any number of syllables, and one syllable to be associated with any number of tones. Trager and Smith's transcription system can represent the first phenomenon, but it is embarrassed

by the association of several tones with a single syllable. They can transcribe a syllable associated with two tones, provided it is phrase-final, by writing one number before the syllable and the second after it. They can handle three tones by adding a rising or falling terminal juncture after the syllable. Four tones, however, is a problem. Trager (1964) makes it clear that this is *not* an oversight. If Trager and Smith's description were a theory, it would predict that no syllable is associated with more than three tones (or two tones and a terminal juncture). This is incorrect. A monosyllabic phrase with the Switchback morpheme will be attached to four tones: 1313. It looks like this:



11. 'four

1313

See Chapter 3 for discussion of this tonal morpheme.

2.2. Terminal junctures

How do Trager and Smith manage with "only" four tones, where we require five? There is a simple answer to this question. I do not find this answer entirely satisfactory, convincing though it seems, but I will leave my doubts until later. Trager and Smith need only four pitches because they have terminal junctures as well as pitch phonemes. They make it perfectly clear (1951:45) that not using terminal junctures would necessitate "setting up, at a minimum, an extra phoneme */5/".

2.2.1. We use the fifth tone to account for the rise in pitch at the end of a High Rise phrase. Trager and Smith would account for this by means of a rising terminal contour, which "is the principal phonetic characteristic of DOUBLE-BAR juncture, /|||/" (1951:46), as in the following.

⁴hów + do + thèy + stûdy | | (p. 51)

'how do 'they 'study

4

5

The appearance of terminal junctures in intonation patterns introduces a difficulty which Trager and Smith ignore. Their intonation patterns are odd-looking morphemes because they contain two kinds of phonological units: pitch phonemes and terminal junctures. Our tonal morphemes, by contrast, contain only tones. Like nontonal morphemes, which contain only phonemes, they consist entirely of phonological units of the same kind. If we adopt Trager and Smith's analysis, we complicate not only our representation of tonal morphemes, but our theory of English as a whole. We would lose an attractive generalization about the structure of morphemes. In fact, this generalization may hold for all languages. It might be claimed Chinese words falsify it, since simple Chinese lexemes have traditionally been described as morphemes consisting of one tone and at least one phoneme, but in an autosegmental framework these words can be analyzed as two morphemes, one tonal and one nontonal. These morphemes would not have meanings, but this is not surprising. English morphemes like *cran* and *cept* demonstrate that morphemes do not necessarily have meanings.

2.2.2. If Trager and Smith's analysis of intonation patterns is valid, then an attractive hypothesis about English linguistic structure is incorrect. Since the consequences of accepting terminal junctures are far-reaching, we will need very good reasons to adopt this solution. Because Trager and Smith had considered some of the alternatives to terminal junctures, they explain the reasons for their choice. Let us see whether they are strong enough to justify the abandonment of our hypothesis about the structure of morphemes. They cite "the following facts" as crucial to their decision (1951:45-6):

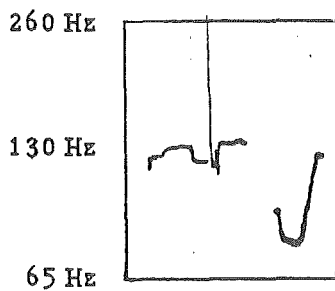
- (a) the exact pitch involved depends on the allophone of the pitch phoneme preceding the contour as a starting point;
- (b) any rise reaches a point well below the next higher pitch (if there is one), and the absolute height reached is a function of the starting point;
- (c) any sustention maintains pitch at the starting point until terminal silence is

reached;

(d) any fall quickly moves down to silence.

Only the first two of these points requires comment. Regarding (a), since the height of a pitch phoneme varies directly with stress, they are claiming that the end-point of the rise is determined by the stress phoneme on the final syllable. That is, the more prominent the syllable, the higher the final pitch. I do not believe that Trager and Smith had the equipment to test such a claim, much less establish it as a fact. It is a testable claim, but not worth testing, since its truth depends on the correctness of their claim that the absolute height of a pitch phoneme is a function of stress, and we have already seen that this hypothesis is incorrect.

With regard to point (b), it is certainly not a fact that any rise reaches a point well below the next higher pitch. Consider the trace below, which belongs to a Fall-Rise. In our theory the beginning point is two steps above the lowest point, but for the sake of argument let us say that they are only one level apart. The end point is clearly not "well below the next higher pitch".



12. 'can I 'help you

3 3 1 3

2.2.3. I think it is fair to say that Trager and Smith's arguments are not particularly cogent. It is not really clear why they prefer to add three phonological entities of a quite different kind rather than increase the number of pitch levels by one or two, but we can point to some considerations which may have led them to promote the complicated solution above the simple one. I will briefly discuss three factors.

First of all, what we might call dynamic phonological entities were part of their linguistic inheritance. All of the five pitch phonemes identified by

Bloomfield (1933:92) are dynamic. They are certainly not pitch levels:

- [.] . . . the falling pitch at the end of a statement;
- [¿] . . . the rising-falling pitch at the end of a question to be answered by speech forms other than *yes* or *no*;
- [?] . . . the rising pitch at the end of a yes-or-no question;
- [!] . . . the distortion of a pitch-scheme in exclamations;
- [,] . . . the pause, often preceded by rising pitch, that promises continuation of the sentence.

Secondly, Trager and Smith are already committed to the kind of phonological unit which they call a juncture phoneme. Their internal juncture, /+/, occurs between phonological words. They use it to account for differences in the pronunciation of particular segmental phonemes which depend on their position in a word, e.g. final stops are unaspirated. Other linguists account for this phenomenon by having phonological rules which refer to syntactic and morphological structure. This solution is not available to Trager and Smith because of their positivist approach to linguistics. They want the environments in which the different allophones occur to be described only in terms of phonetic units. This leads them to claim that there is a phonological unit called an internal juncture, which has a wide range of allophones manifested in the surrounding phones. But the point here is really that the "discovery" of terminal junctures provides support for an analysis using internal junctures, since the latter are no longer one of a kind-- or so Trager and Smith might argue.

Thirdly, we have seen that the postnuclear rises in our English tonal morphemes always occur on the final syllable of the phrase. Trager and Smith evidently observed this as well, and wanted their system to capture this generalization. This sort of thing is not really part of their brief as structuralist phonologists. There is no difficulty in describing or transcribing all the relevant phonemic clauses in terms of pitch levels, though one might use five rather than four. For a structuralist linguist using pitch phonemes, the fact that rises are always final is not a matter for phonology, but rather for morphology. In our theory we are able to say that all of these final tones are attached by a particular kind of initial tone association rule, one that links a particular tone

to the final syllable of a phrase. Since Trager and Smith do not have tone association rules or their equivalent, this kind of explanation is not available to them. They must resort to a new kind of phonological unit-- the rising contour, in the guise of a terminal juncture.

2.3. Redundant and missing pitches

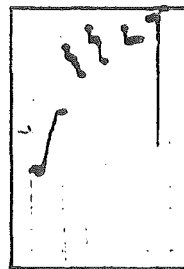
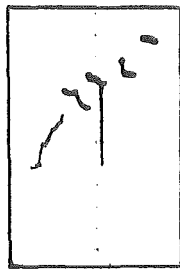
We have now looked at some possible explanations for Trager and Smith's introduction of terminal junctures in spite of the fact that they are an unnecessary complication. Prior to that I argued that terminal junctures allow Trager and Smith to describe English tonal morphemes using only four tones, or pitch levels. This, however, is an oversimplification.

Trager and Smith's pitches 3 and 4 do not in fact line up with our tones 3 and 4. Consider the following pair of traces, for example. I have spoken these to illustrate Trager and Smith's transcriptions (pp. 50-1).

260 Hz

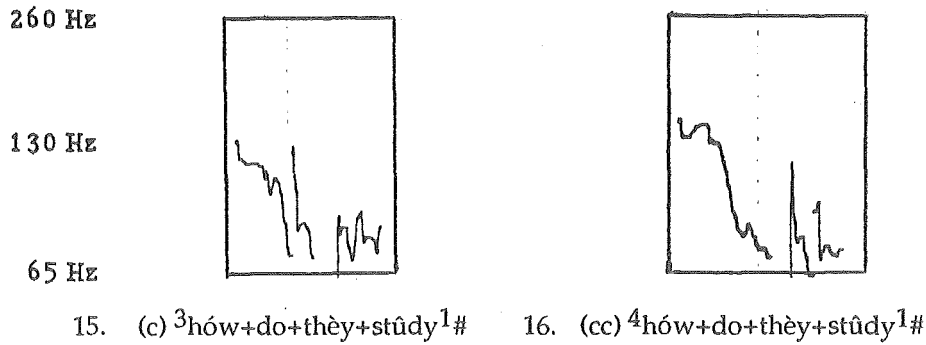
130 Hz

65 Hz



13. (e) ³hów+do+thèy+stûdy || 14. (ee) ⁴hów+do+thèy+stûdy ||

Trager and Smith would say that the second phrase begins with a different pitch phoneme from the first-- 4 rather than 3. We would say that both phrases begin with the same tone-- 4, in this case-- but that the actual pitch of this phonological unit is higher in the second phrase. There are two assumptions behind our claim. One is that native speakers of the dialect in question would identify these phrases as carrying the same tune. The other is that each tune has only one underlying form. Consider another pair:



Again the difference according to Trager and Smith is that the second phrase has a 4 pitch where the first has a 3. In our representation the first tone would be the same-- 3, probably-- in each. The higher pitch of the tone in the second phrase would be the result of the operation of later, or more superficial, phonological/phonetic rules.

From our point of view, therefore, the fourth pitch phoneme of Trager and Smith is redundant, an emotional variant of the 3 pitch. There is support for this view in their own observation that "In any of the illustrations having /3/ before /1/, /11/, or /#/ , we may substitute /4/ and get exactly parallel results" (1951:46). This argument appears to lead to the conclusion that we do not need five tones. If Trager and Smith are not exploiting their fourth tone, then four tones should suffice for a theory without terminal contours, like ours. But this is not the whole story, because Trager and Smith are really one tone/pitch short.

In all the common standard English tonal morphemes, the prenuclear accent tone is a 3-tone. The pitch which Trager and Smith use in this position is a 2-pitch. We do not use a 2-tone here because the nucleus tone of the Terrace tune is a 2-tone, and it is lower than the standard accent tone. I cannot find among Trager and Smith's examples any phrases with the Terrace tune. This is not surprising, because they have only a few examples. Clearly they would transcribe such a phrase with a /1/ terminal juncture-- sustention-- and they would probably use the "correct" tones-- 3 and 2. The problem is that they would transcribe the prenuclear part of a Low Bounce-- and other tunes-- with a 2-pitch, in spite of the fact that it is the same as the prenuclear part of a Terrace.

3.2. Wells

Wells's (1945) theory is, like Pike's, a pure levels theory. He does without contrastive terminal contours. As with Pike, the key to his avoidance of the fifth tone is the High Rise. Unlike Pike, he represents the nuclear syllable in High Rise phrases as one step higher than a preceding accented syllable. In his example no. 24, the first syllable, which I take to be accented, is associated with a 3-pitch, while the second syllable, which I take to be accented and nuclear, is associated with a 4-pitch.

(24.) ³*Give* ⁴*him the money?*

All of the syllables following the nuclear syllable are associated with the same pitch phoneme as the nuclear syllable. In other words, the transcription does not indicate the rise in pitch on the final syllable which is characteristic of this tune. This is true also of his other examples of this tune:

(27.) ⁴*Who did you say did it?* (great surprise)

(1.) ²*Are you* ³*positive?* (Also: ²*Are you* ⁴*positive?*)

Wells does recognize the existence of this rise, but he makes it clear that he does not wish "to set up a pitch phoneme 5 just to account for this rise" (p. 33). Instead he treats the pitch of the final syllable as a high allophone of 4, apparently on the grounds that 2 and 3 have similar allophones. Apparently Wells believes that the final rise is non-distinctive, i.e. that there is no contrast between a High Rise tune ending with 4⁵ and a High Level tune ending with 4.

It is not difficult to guess why Wells misses the High Level tune. It is characteristic of nonfinal phonological phrases, phrases which in standard orthography would be followed by a comma or a dash, if anything. Each of Wells's examples, on the other hand, is complete-- "a single clause", as he puts it (p. 32)-- and ends with a question mark, an exclamation mark, or a full stop.

In order to provide support for my introduction of a fifth tone, I have looked at 4-tone analyses to see how they fail. We have seen that some of these analysts actually mention the possibility of a fifth tone, but are reluctant to introduce it because it occurs in only one tonal morpheme. In segmental phonology and morphology such a consideration might carry some weight, because the number of nontonal morphemes in English, and in languages in

general, is relatively large. A phoneme which appeared in only one of 100,000 morphemes would certainly be suspect. Tonal morphemes, however, are much fewer in number. A tone which appears in only one of ten morphemes is not suspect at all. In tonal languages it is not unusual for a particular tone to occur in only one tonal morpheme. For example, in the Black Miao language described briefly by Anderson (1978:145), tones 2 and 4 each occur in only one of the eight tonal morphemes.

The findings of students of tone languages also support my conjecture that English uses five tones. Anderson (1978:146) asks how many (level) tones a universal grammar would need to provide, and concludes that "a phonological feature system must provide for the description of at least (and apparently at most) five levels of tone". He cites data from a Chinese language and an African one which clearly contrast five tone levels, and refers to other studies where the same claim is made. The work of Gandour (1978) and others on tone perception also supports the recognition of five phonological levels of pitch. Of course, it does not follow from this that English must have five tones, since by no means all languages exploit all the possibilities made available by the universal grammar. It does mean that the extra pitch which Trager and Smith and Wells are reluctant to postulate "costs nothing". The conclusion of the tone language specialists also provides support for my claim that no more than five tones are required, but this does not seem to be in dispute.

Chapter 5

Postnuclear Accent Tones

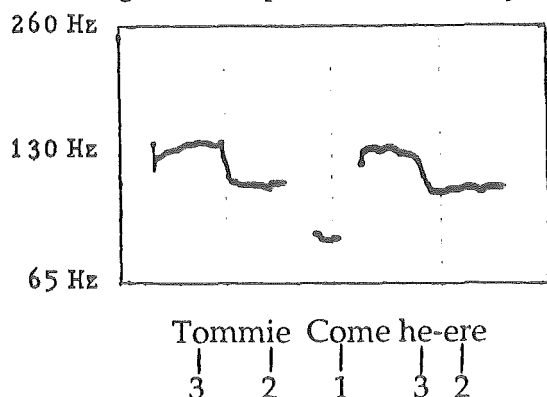
The hypothesis under attack in this chapter is that all postnuclear tones of an English tonal morpheme are either unmarked or end tones. I will argue that the morpheme known as the calling tune contains an accent tone which follows the nucleus tone. This has several consequences. The most significant is that it throws into question the hypothesis that the nuclear syllable is the last accented syllable in a phonological phrase. It also suggests that correct operation of the accentuation rules requires information about the tune attached to a phrase. Evidence to support the latter hypothesis, which affects the position in a derivation of tonal morpheme insertion rules, comes from analysis of a tune used by stock auctioneers.

The tonal morpheme whose representation concerns us here has been labelled in various ways, but most of them reflect the observation that its most easily identifiable use is for calling or hailing. The earliest description seems to be that of Pike (1945:71), who includes it in a rather small class of "spoken chants". Pike transcribes two examples as follows. The colon indicates extra length.

Tommie! come here!

°2: - 3:/ 4: - °2:-3:/

Below I give the equivalent in our system of representation.



More recently the tune has been discussed by Fox (1969, 1970), Crystal (1969b), Lewis (1970), Gibbon (1976), Leben (1976), Liberman (1978), and Ladd (1980). Pike identifies the constituent pitches as 423, which corresponds to LHM in some theories and 132 in ours, and this is widely accepted. Pike's transcription also makes it clear that the first tone is optional, and that the second tone is the nucleus tone. Again, there is little disagreement with these claims.

Most investigators would also agree that this tonal morpheme has at least two unusual characteristics. The first is that each of the obligatory tones demands its own syllable. If the nuclear syllable is final, or as Fox (1969:13) puts it, "if there is no tail, the nucleus will be split into two". This is illustrated by the word *here* in the examples above. Thus the calling tune falsifies the hypothesis that the obligatory tones of a tune can be attached to any number of syllables. This seems to be one of the characteristics which distinguish Pike's spoken chants from other tunes.

The second unusual characteristic, which has to do with the final tone, is more exceptional. In all other tonal morphemes we have examined, a postnuclear tone is attached either to the last syllable, or to the first postnuclear syllable and any following it. The final tone of the Dipper morpheme, (1)(3)313, illustrates the first possibility, and the penultimate tone illustrates the second possibility. The final tone of the calling tune poses a problem because it is like neither of these. If there is a strong syllable after the nuclear syllable of the phrase, as in 'Marma'dukie and 'Lunch is 'ready, the 2-tone will be attached to it and to any following tones.



If there is no strong syllable between the nuclear syllable and the final syllable, as in 'Pamela and 'Jefferson, then the 2-tone is attached only to the final syllable.



Of the linguists mentioned above, only Leben and Liberman, who seems to follow Leben in this matter, appear to have addressed the problems which

this aspect of the calling tune presents for a theory of English intonation. Before explaining and defending my solution in section 2, I will outline that of Leben, and discuss some of its inadequacies. In section 3, I will consider some of the consequences of adopting my representation.

1. Tone Spreading Solution

Leben's approach is to make the underlying representation of the tune unexceptional, and add a tone rule which applies relatively late in the derivation to modify the output of the ITA and well-formedness rules, which would otherwise be wrong.

1.1. How it works

Leben (1976:98) summarizes the relevant data as follows. (His starred syllable is our nuclear syllable, and his L, M, and H, are the equivalent of our 1, 2, and 3-tones, respectively.)

- i) Before the starred syllable, all syllables are L.
- ii) The starred syllable is H.
- iii) If following the starred syllable there are only reduced vowels, then all are H except for the last, which is M. If the starred syllable is also the final syllable, then it is drawn out into two syllables, the second of which has M.

*

Mister Smi-ith
L L H M

- iv) If there is at least one unreduced vowel after the starred syllable, then all vowels after the starred syllable are H up to but not including the first unreduced vowel. The remaining vowels are all M.

Leben proposes to account for these facts by representing the tune as LHM with the H starred-- i.e. 132. Since the final tone is not marked, it will not be attached by an initial tone association rule. The well-formedness rules will then attach the 2-tone to the immediately postnuclear syllable and also to any following syllables. As Leben makes clear, this results in incorrect representations, such as the following. (Bold type will henceforth replace Leben's star on nuclear syllables.)

*Pamela
 3 2

*Pamela dèar
 3 2

*Mister Jefferson
 1 3 2

"To convert these representations into correct ones," says Leben (1976:101), "a rule must be formulated to propagate H rightward, stopping just before an unreduced vowel or just before the final syllable, whichever comes first. The symbol ə will denote any reduced vowel, and X and Y stand for the surrounding segments."

TONE SPREADING (OVER MID)

X ə Y ==> X ə Y
 | | |
 H M H M

e.g. Pamela dèar ==> Pamela dèar
 | | |
 H M H M

i.e. 'Pamela 'dear ==> 'Pamela 'dear
 3 2 3 2

1.2. Weaknesses

In this subsection I discuss the weaknesses of Leben's solution, starting in 1.2.1 with those which he mentions in his paper, and then turning in 1.2.2 to some he does not mention.

1.2.1. Leben describes two objections to a rule of tone spreading (101):

The first is that it is ad hoc, since it seems designed solely to make the strong autosegmental treatment consistent with observational adequacy. [That is, a single tune would otherwise have to have more than one underlying form--DH. "The strong autosegmental hypothesis forbids representing a single contour by two formulas, with the choice between them governed by segmental differences" (p.99)] The second objection is that if English had only the Tone Spreading rule which extends the domain of a H tone over M, it would violate an implicational universal formulated by Hyman and Schuh [1974] whereby languages exhibiting spreading of H over M must also exhibit spreading of H over L.

He attempts to answer both objections by claiming that "a number of other cases of Tone Spreading are attested in English, some of them applying to the sequence HL". We shall see that the case Leben considers most convincing fails to support his argument, because the tune does not behave as his hypothesis predicts. He refers to an "intonation pattern indicating condescension or mockery", and gives the following examples, among others. (Instead using association lines, Leben writes a tone symbol under each syllable. When none of the syllables in a diagrammed phrase is associated with more than one syllable, I will sometimes follow Leben in omitting all the tone association lines, but I will place each tone number only under the leftmost syllable to which it is attached.)

So you're a Republican!

H H L

I see your husband's name is Pamela!

H H L

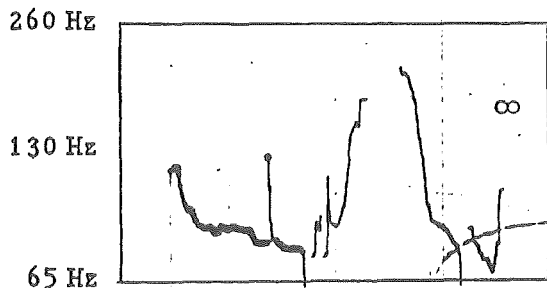
So you plan to elope with our secretàry!

H H L L

So you wanna be a bank prèssident!

H L L L

Leben does not supply any phonetic data for these. The following trace is my attempt to illustrate the last of his examples.

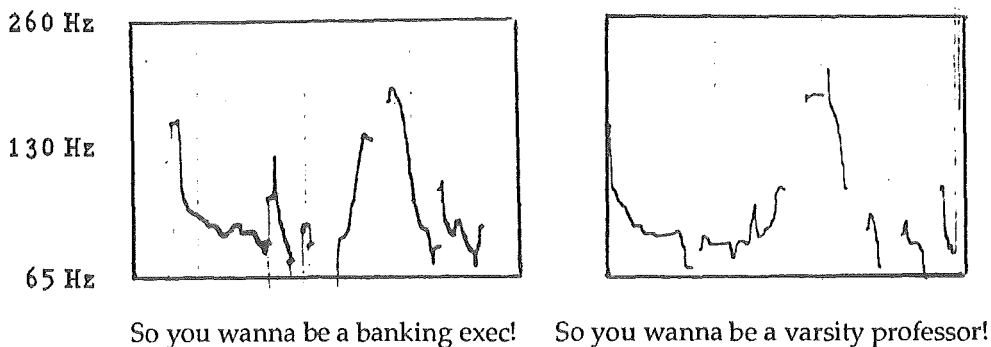


So you wanna be a bank prèssident!

According to Leben (p. 101) these differ from "the corresponding sentences having neutral declarative intonation" in that the fall from H to L is delayed.

He seems to be describing what other writers call the rise-fall tone. His subsequent description of this pattern as a "sarcastic chant" (p. 105) supports this identification, since Cruttenden (1986:102) says that a "local meaning sometimes ascribed to the rise-fall is 'ironic' or 'sarcastic'." Leben goes on to claim that "the condition determining where the fall is placed is the same as in the vocative chant: H moves rightward over as many unstressed syllables as it finds; if an unreduced vowel is encountered, as in *secretary*, *bank* *président*, the L begins on this vowel; otherwise, the L occupies only the final syllable, as in *Republican*, *Pamela*".

Leben makes no attempt to test this hypothesis. It predicts that any and all weak syllables between the nuclear syllable and the first strong postnuclear syllable will be associated with a High tone. None of his examples, however, contains more than one weak syllable in this position. This is not exactly pushing a hypothesis to the limit. What does in fact happen when there are two or three such syllables? We can create examples by replacing '*bank* *'président*' with '*banking* *e'xecutive*' (or *exec*, to simplify interpretation of the trace), which has two weak syllables, and with '*varsity* *pro'fessor*', which has three weak syllables before the strong *fess*. The traces below show that none of the weak syllables following the postnuclear syllable is at all high. Since Leben's rule of Tone Spreading would attach all of them to a High tone, his hypothesis is clearly wrong. It is to Leben's credit that his hypothesis is clear and precise enough to test easily.



Leben discusses other purported examples of his Tone Spreading rule at work, but he himself recognizes that they are less convincing, and relies on the sarcastic chant to support the rule (105). Since Tone Spreading does not in

fact work for this chant, this support vanishes.

1.2.2. Another weakness of the tone spreading solution to the calling tune has to do with the contours which are distinguished by the operation of the rule Leben proposes.

If Tone Spreading applies to the sequence HL, the result is the "sarcastic chant", or rise-fall. If it does not apply, the result is "neutral declarative intonation", or fall. These two intonations would therefore have the same underlying form. That is, they are manifestations of the same tonal morpheme. In my theory, too, the rise-fall would be derived from the (1)(3)31 morpheme, by the application of boosting and stretching rules. Both are independently motivated, and both are associated with emphatic or enthusiastic speech. If Leben's rule of tone spreading worked, it would be an appropriate way of accounting for at least some of the differences between the neutral declarative and the sarcastic chant. It is less appropriate, however, for the calling tune.

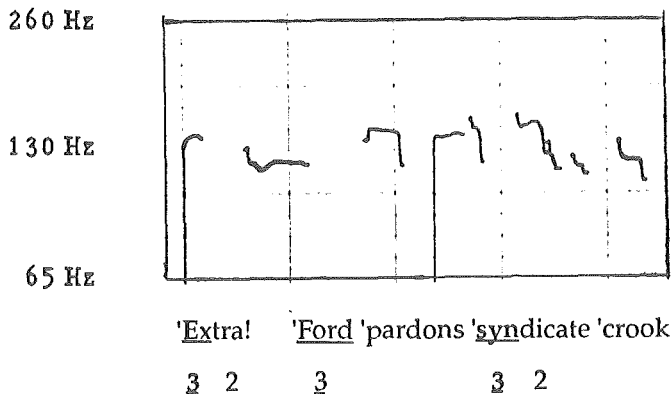
If Tone Spreading applies to the sequence HM, the result is the "vocative chant", or calling tune. If it does not apply, the result is another chant, "the one typically employed by newsboys and train conductors" (97). Leben illustrates this as follows:

Extra! Ford pardons syndicate cròok!

H M H H H H M M M

Next stop Ottawa!

H H H M M



In chants of this type, the M (or 2) tone is attached to all the syllables which follow the nuclear syllable. According to Leben, then, the newspaper vendor chant and the vocative chant have the same underlying form. It does not seem right, however, to treat these two chants as manifestations of the same tonal morpheme. Leben himself says that if the drop in pitch is delayed until the final syllable of the two examples above, the resultant phrases "sound distinctly peculiar in this context" (p. 97). This bizarreness is not surprising if we are dealing with two different tonal morphemes. Replacing a vocative chant with a vendor chant results in the same bizarreness or unacceptability. This is not what happens, though, when we replace a sarcastic chant with the tune which, according to Leben, is related to it as the vendor chant is related to the vocative chant. When we replace a sarcastic chant with "neutral declarative intonation", in the same context, the results do not sound peculiar.

Summing up, we have seen that Leben's derivation of the calling tune by means of his Tone Spreading rule has at least two related weaknesses which he does not discuss. First, it claims that the calling tune differs from the vendor/conductor tune in the same way that the sarcastic chant or rise-fall tune differs from the plain falling tune. Second, it claims that the calling tune has the same underlying form as the vendor/conductor tune.

2. Accent Tone Solution

Instead of accounting for the difference between the calling and conductor tunes by means of a Tone Spreading rule, let us say that the 2-tone of the calling tune is attached by an initial tone association rule, while the 2-tone of the conductor chant is not. If we represent the latter as (1)(3)32, the operation of the well-formedness rules in the usual way will produce the desired results, assuming that we have a syllable splitting mechanism which will come into effect when there are no postnuclear syllables.

We now face the problem that none of the ITA rules which we have been using will attach the 2-tone to the correct syllable. We have assumed, following Clements and Ford (1979:181), that there are three types of ITA rule

available:

1. Associate a designated tone of the tonal string with the tone-bearing unit that bears the accent.
2. Associate a designated tone of the tonal string with the leftmost (or rightmost, penultimate, etc.) tone-bearing unit.
3. Associate tone n of the tonal string with tone-bearing unit m (counting from the left). (Usually $n = m = 1$.)

Two ways around this difficulty suggest themselves. One is to change the ITA rule schemata, either by adding a new one or by modifying one of these, probably the first. The other approach is to change the accentuation rules so that the first kind of rule can be used as it stands.

2.1. First alternative

There are at least three reasons for not following the first course. The first is that the new ITA rule would have to be considerably more complicated, or powerful, than the existing ones. Assuming that it comes into effect after the nucleus tone is attached, it will be required to determine whether there are any available strong syllables. If there is at least one, the tone will be attached to the leftmost of them. If there are no strong syllables, the tone will be attached to the final syllable. The extant ITA rules have only to locate a particular syllable. They involve no *ifs* or *buts*. The contrast between their character and that of the proposed new rule suggests that this is not the most natural solution.

A second reason is that we will have to change the accentuation rules even if we do modify the ITA rules so that they will attach tones to unaccented syllables. When the 2-tone is attached by default to the final syllable, as in *Pamela*, this syllable is perceived as strong. As Fox (1969:13) puts it, "if no prominent syllable follows, a normally weak syllable will be made strong for the purpose". Since tone association rules cannot change the value of the feature [strength], this will need to be done by a new or modified accentuation rule.

A third reason is that these ITA rules were not hypothesized for English. They are supposed to be universal, and they are apparently able to account for

data from a variety of languages. It would be ill-advised to tamper with them in order to handle a single English tune whose form and status are both debateable. Simple hypotheses about universal grammar ought not to be abandoned lightly. Let us therefore suppose that the ITA rules are correct, and seek a solution which incorporates them as they stand.

2.2. Second alternative

It is clear that the third type of ITA rule does not apply. The second type appears to apply if there is no postnuclear strong syllable, since the tone is then attached to the final syllable, but it would have to be used in conjunction with another rule which operates when there is a strong syllable. This is unprecedented. It would be unwise, I think, to abandon at this point the assumption that only one ITA rule, at most, may apply to a particular tone. We are left with the first type of ITA rule, which will work if we suppose that the 2-tone of the calling tune is an accent tone. If it is, then phrases attached to this tune must have an accented syllable following the nuclear syllable. If we permit this, we must abandon the hypothesis that the nuclear syllable is always the last accented syllable in a phonological phrase.

In particular, we will need to modify the accentuation rules, apparently by adding a postnuclear accentuation rule. This rule will scan from left to right the postnuclear syllables of a phrase, and make accented the first [+strength] syllable it finds. If it reaches the end of the phrase without finding such a syllable, it makes the final syllable accented. This solution has been more or less forced on us by our decision not to tamper with the ITA rules, but there are a number of facts which suggest we are on the right track.

2.2.1. First of all, this new rule looks familiar. It does the sort of thing which accentuation rules do, such as locating the first stressed syllable in a sequence. In this it is unlike the proposed ITA rule discussed above, which was different in character from established ITA rules. In fact, this postnuclear accentuation rule operates just like the prenuclear accentuation rule proposed in Chapter 2, which makes the first [+strength] syllable in a phrase accented. The new rule does the same thing in a different place. Its domain is the postnuclear part of

a phrase rather than the prenuclear part.

It appears therefore that we do not need a completely new accentuation rule for the sake of the calling tune. A rule of this kind is independently motivated. We need only a single nonnuclear accentuation rule, which will operate in more than one domain. This rule must have a default clause which operates when the domain contains no stressed syllable. In this case, the final tone of the calling tune is attached to the final syllable, so the accentuation rule must make the final syllable accented. This does not happen when the nonnuclear accentuation rule applies to the prenuclear part of a phonological phrase. The obvious explanation for this is that postnuclear accent tones are obligatory, while prenuclear accent tones are not, as we have been assuming. We can only use this as an explanation, however, if we assume that accentuation rules are supplied with information about the form of the tonal morpheme to be attached by the tone association rules. I shall return to this problem below.

2.2.2. We move on to a second kind of support for the hypothesis that the final tone of the calling tune is an accent tone. This hypothesis forces us to predict that the first postnuclear syllable after the drop in pitch, the first syllable attached to the 2-tone, is accented. Recall that accented syllables are more prominent than strong syllables which are not accented. Nonnuclear accented syllables, since they share all but one feature specification with nuclear accented syllables, may be difficult to distinguish from nuclear syllables. Our prediction, unlikely as it seems initially, may well be correct.

First of all, our hypothesis explains why some students of intonation have identified such syllables as nuclear. This is not surprising if they are accented. Crystal (1969b: 35) explicitly claims that the first mid-pitched syllable in phrases of this kind is nuclear, carrying a level nuclear tone. He chooses to "analyse *Johnny* (or the *Jo-ohn* type, which is of course phonologically disyllabic) as a case of a final level tone preceded by an extra-prominent syllable" (36). Other analysts are less explicit, but they, too, describe the syllable in question as nuclear, and therefore, according to our definition, accented as well. For instance, some of O'Connor and Arnold's examples of their Terrace

tune appear to illustrate the calling tune. When this tone group is used in final word groups like the following, they say (1973: 89), "it gives the impression of calling out to someone, as if at a distance".

'Dinner's >ready. 'Thank >you.

The marks preceding the syllables *rea* and *you* indicate that the authors consider them to be nuclear, and that they are lower than the syllables preceding them. According to our hypothesis, these phrases would be represented as follows.

'Dinner's ready 'thank 'you
 3 2 3 2

Our hypothesis that the first syllable attached to the 2-tone is accented is also supported by what Crystal (1969b:34) describes as "the potential of either element to be increased in duration for as long as one likes". In other words, either the nuclear syllable or the postnuclear accented syllable, or both, can be stretched. Others have made this observation, too. Devising a rule for this is straightforward if only these two syllables share the feature specification [+accent].

If the syllable in question is accented it should be more prominent than a syllable which is strong but not accented. Thus, in the following example, which I overheard called across a busy street, our hypothesis predicts that native speakers will think *look* "sounds louder" than *for*, even if the vowel of the latter has its full quality and length. This appears to be correct.

Your 'mother's 'looking 'for you
 1 3 2

2.2.3. Let us turn to a third kind of evidence which could support the notion of a postnuclear accent tone in the calling tune. We have been assuming that every English tonal morpheme contains one and only one accent tone as well as one nucleus tone. In the morphemes discussed so far, the accent tone is prenuclear and optional. If this hypothesis is correct, then a tonal morpheme which contains a postnuclear accent tone will not contain a prenuclear accent tone. If we find that the calling tune lacks a prenuclear accent tone, this will support our hypothesized postnuclear accent tone. According to some

descriptions of this morpheme, it does indeed lack a prenuclear accent tone. Fox (1969:13), for instance, says that "the pretonic is low to mid and is level: *come* and *get -it*, *where* *are -you*". That is, all the prenuclear syllables are relatively low, attached to a 1-tone in our terminology, whether they are strong or not. If this is so, we can represent the tonal morpheme as (1)32. Liberman (1978:20), too, states that "If there are any syllables preceding the main stress, the low tone is associated with them", and his examples include the name *Tippecanoe*, which has a stressed initial syllable. The position of other investigators on this question is not always clear, but agreement is not universal. Leben cites the following examples, apparently supplied by Bolinger:

Colette and Jimmy!	Everybody out of the poo-ol!
L H H H M	H H H H H H H M

If we wish to modify our representation in order to account for these examples, we will have to insert an optional accent 3-tone before the nucleus tone: (1)(3)32. It seems therefore that we cannot expect support for our final accent tone from this source. We will apparently have to abandon the single accent tone hypothesis mentioned above.

3. Tune-Sensitive Accentuation Rules

2.3.1. Introduction

We have now seen that, even if the calling tune does contain a prenuclear accent tone, there is considerable support for the hypothesis that the postnuclear 2-tone is an accent tone. Let us now return to the postnuclear application of the nonnuclear accentuation rule, which is necessitated by our analysis of the calling tune. This rule is problematical in that its application must be restricted. The ideal accentuation rule applies to all phonological phrases. If we allowed this one to apply to all phrases, we would in effect be making several claims which are demonstrably incorrect. These include the following three: (i) if there is more than one strong syllable following the nuclear syllable, as in *En'tirely incor'rect*, the first of them is always accented and therefore more prominent than the others; (ii) if there are no strong

syllables between the nuclear syllable and the final syllable, as in '*Long phrases*', the latter must be accented, and therefore as prominent as a preceding accented syllable; and (iii) if there is a weak syllable between the nuclear syllable and the final syllable, as in '*Syllables*', it is always less prominent than the final syllable.

Let us therefore suppose that the rule in question applies only to phrases which contain or carry this particular tonal morpheme. This is a new notion. One of our basic premisses is that the accentuation rules have operated before the tone association rules link the syllables of a phrase with the tones of a tonal morpheme. This assumption is safe, but we cannot say the same about the assumption-- unspoken up to this point-- that accentuation rules do not require information about the identity or form of the tonal morpheme to be attached by the tone association rules. It now appears that accentuation rules must have access to such information. If this is so, whatever rules select the tonal morphemes do not operate immediately before the tone association rules. The identity of the morpheme must be available to rules earlier in the derivation. It is not just the postnuclear accent tone which makes this necessary. It would be necessary anyway for the rule which splits a phrase-final nuclear syllable if the phrase carries the calling tune or the conductor tune. The simplest solution is to have this rule operating before the accentuation rules. The rules we have proposed will then produce the desired results. The new postnuclear syllable will be accented by the postnuclear accentuation rule, and it will be linked to the 2-tone by the applicable ITA rule. There seems no reason to introduce further complications by locating the rule later.

I have argued that the rule which accents postnuclear syllables is in essence the same as the one which accents prenuclear syllables. It cannot be denied, however, that this particular application of the rule is ad hoc. It explains the data it was designed to explain, and nothing else, apparently. There are two defences, or rather justifications, for this approach.

First of all, having a rule for only one tune amounts to a claim that the tune is exceptional. For the calling tune, this claim appears to be correct. There is general agreement that it is unusual in other respects as well. Pike

(1945:71), for instance, includes it in a very small class of tunes, spoken chants, whose rhythm is syllable-timed rather than stress-timed. Also exceptional is the tune's need for two syllables. So is the characteristic which Gibbon (1976:274) labels "chroma".

Secondly, this tune is not unique in requiring a special accentuation rule. In particular, such a rule is also required for a tune, the presale tune, used by stock auctioneers. To demonstrate this I will need to go beyond my own dialect to the language of stock auctioneers in New Zealand, especially the Christchurch region. This tune may well occur in British and Canadian stock auctions, but my investigation of the NZ variety has been more detailed.

3.2. Stock auction presale tune

The discourse structure, formulaic syntax, and prosody of New Zealand stock auction speech are described in Kuiper and Haggio 1984 and, more briefly, in Haggio and Kuiper 1985. In many of the auctions they report on, there occurs a tune which they call the End Tune. They argue that this tune functions as a prosodic signal that the bidding is about to end. That is, it functions in the same way as what they call presale formulas, which act as verbal signals of the approaching sale. These include *on the market, here for sale, we'll cash them, I'll sell 'em sir*, and *gonna sell 'em now*. The End Tune, or presale tune as I call it in Haggio 1983, appears to contain the same tones as the calling tune, in the same order: 132. The words of the phrase most often represent an amount, the amount of the previous bid. (The letters and numbers in parentheses identify the auctioneer and the auction.)

twenty dollars forty (PMD A2)

1 3 2

twenty three dollars sixty (DFF A10)

1 3 2

twenty one dollars (PAH A25)

1 3 2

twenty three twenty (FT A34)

1 3 2

All three tones appear to be obligatory. The last of these, the 2-tone, is the nucleus tone. We know this because familiar tunes are widely used elsewhere in the bidcalling, and the nuclear syllable is consistently in the last word. All the following examples come from the auction transcribed in Appendix 1 of Kuiper and Haggo 1984 (DFF A10). I omit irrelevant details here.

twenty two sixty

3 1

at twenty two dollars sixty

1 3

three dollars

3 1

at twenty three sixty

1 3

The 3-tone is more interesting. In all the examples above the first high pitched syllable is the third syllable of the phrase. If this were true of all phrases bearing this tune, the 3-tone could be attached by an initial tone association rule of the third kind, which we have not used yet. Such a rule links the n th tone with the m th syllable of a phrase; n would be 2, and m would be 3. The discovery that English used such a rule would be big news. In fact it does not. If the first word is trisyllabic, like *seventeen*, it is the fourth syllable of the phrase which is the first high one:

seventeen dollars fifty (DFF A14)

1 3 2

What we need, apparently, is an ITA rule which will attach the 3-tone to the first syllable of the second word. None of our ITA rules will do this. In all of these words, however, the first syllable is also the first underlyingly strong syllable. Our prenuclear accentuation rule operates on the first strong syllable

of specified words. Let us suppose that the 3-tone is an accent tone, and therefore attached by an ITA rule of the first kind. We therefore propose that the first stressed syllable of the second word of these phrases is accented. The normal accentuation rules will not, however, produce the results we want.

In Chapter 2 I discussed two kinds of prenuclear accentuation rule. One accents the first [+strength] syllable of the first [+focus] word of the phrase; the other accents the first [+strength] syllable of the first content word. Assuming that focus is related to importance or informativeness in a particular context, neither type of rule will produce the following:

'twenty 'dollars 'forty

'seven'teen 'dollars 'fifty

The second word, *dollars*, is clearly less informative than the first. Indeed, it is often omitted, in auction speech as in ordinary English. The obvious solution is to introduce an accentuation rule which accents the first stressed syllable of the second word. The term *word* predicts a certain amount of variation, since a number like *twenty three* can be interpreted as one word or two. Such variation does in fact occur. In the relevant examples quoted above, *twenty three* is taken as two words, but in the following phrase the number is taken as one word, so the first high syllable is *dol*:

'twenty 'two 'dollars 'sixty 'only (FT A36)

1 3 2

I propose, therefore, that the presale tune be represented as 132, with all three tones obligatory, and that a special rule accents the first stressed syllable of the second word. If these hypotheses are correct, then this tune requires a phrase of at least three words. They would be incorrect if something like the following occurred near the end of the bid calling:

* 'thirty 'dollars

3 2

I have not found any such phrases. Does this mean that the presale tune cannot be used with relatively common bids such as *twenty* or *thirty dollars*? I think not. I would predict that if an auctioneer wanted to use this tune with a two word amount such as *thirty dollars*, he would make it a three word phrase by adding a word like *at* at the beginning, to bear the 1-tone, or *only* at

the end, to bear the 2-tone. Both of the following phrases occur repeatedly, with other tunes, in the auction given in Appendix 2 of Kuiper and Haggo 1984: *at thirty dollars, thirty dollars only*.

3.3. Conclusion

Clearly the application of the accentuation rule just proposed must be restricted. As far as we know, it applies only to phrases which carry the presale tune. The calling tune is therefore not the only one which requires a special accentuation rule. Thus the nature of the presale tune supports our representation of the calling tune. Let us put this another way. Analysis of the calling tune led us to claim that it, unlike other English tonal morphemes, contains a postnuclear accent tone. This required us to propose a postnuclear accentuation rule which apparently applies only to phrases to which the calling tune is attached. This in turn led us to claim that the accentuation rules as a whole must have information about the tonal morpheme to be associated with a phrase in order to produce the accentuation required. This amounts to a prediction that there would be other tonal morphemes which make it necessary for accentuation rules to "know" which tune is involved. Reordering rules for the sake of a single tune is suspect. This prediction proved to be correct, because there is at least one other tonal morpheme, the presale tune used in New Zealand stock auctions, which makes it necessary for accentuation rules to have this information.

It is possible, of course, that changing the focussing mechanisms could give a better account of the data than changing the accentuation rules. I will not pursue this suggestion, however, because it does not appear to affect the point that the calling tune is not unique in necessitating such revisions.

The rule which we have proposed for the presale tune need not be ad hoc, because we can relate it to an exceptional characteristic of that tonal morpheme. I have argued that the first tone of the presale tune, unlike that of other English tonal morphemes, is obligatory. This enables us to make the claim that all tonal morphemes with obligatory initial tones trigger the nonnuclear accentuation rule which we have proposed for the presale tune.

At the moment, of course, we do not know of another English tune whose first tone is obligatory, but our claim is certainly testable in principle. Ad hoc hypotheses, on the other hand, make no testable predictions.

We can make a similar claim about the accentuation rule motivated by the calling tune, since the latter, too, is formally exceptional. Let us hypothesize that all tonal morphemes which contain a postnuclear accent tone trigger the postnuclear application of the accentuation rule discussed earlier. Again, this is a testable claim. One can imagine a tonal morpheme which would falsify it. One such tune would have a postnuclear accent tone which, if there were no stressed postnuclear syllable, was attached to the first postnuclear syllable rather than the last. So far, such a tonal morpheme has not been discovered, but no one has been looking for it.

Chapter 6

Levels versus Contours

In this chapter I defend the hypothesis that English tonal morphemes consist entirely of level tones, and attack the alternative— that some tonal morphemes contain contour tones. The issue is not whether so-called nuclear tones are phonological units or units at all, but rather the correct treatment of contours. Are they to be represented as a sequence of level tones, or pitches, or are they better treated as atomic contour tones? Consider, for example, what I call the obligatory part of the Drop morpheme and others, a falling nuclear tone. It can be represented as the sequence 31, or High-Low. Alternatively, it can be represented by means of a unitary feature such as [Falling], perhaps with another feature indicating its register.

It has become customary to divide theories of intonation up to the 1970s into two classes: levels theories and contours theories. Bolinger (1951), Crystal (1969), Ladd (1980), and Cruttenden (1986) all make this distinction, though Bolinger and Ladd use the term *configuration*, which is probably less confusing, instead of *contour*. As all of these writers point out, the levels approach developed in America, and contours largely in Britain. Our theory belongs with the former. It is essentially a levels theory. In a levels theory, the ultimate constituents of intonational tunes are tones, or pitch-level phonemes, as they are called by some. There are some important differences, of course, between our approach and earlier levels theories. Some of these are taken up in Chapter 4, where I compare this theory with that of Trager and Smith, and account for the differences. The contours approach is not so easy to define positively. I will take the simplest course, which is to define contourists as anti-levelists. That is, those linguists who have been identified with contours

or configurations deny that tunes consist entirely of level tones. Thus most, if not all, of them would agree that atomic contours are necessary for the description of English intonation, though few have said so in these terms. I will refer to this as the English contours hypothesis.

I will begin by attempting to dispel some of the confusion which has characterized the levels versus contours debate. Then I will briefly summarize the most obvious arguments against atomic contours, before going on to consider possible arguments in their favour. Finally, I will examine some of the criticisms which contourists have directed at levels approaches, and show that there is little substance to their critiques.

1. Related hypotheses

I have suggested that the question at the heart of the levels versus contours debate is whether or not English uses atomic contour tones. Reading the literature, however, one would not believe the issue is so simple. This is partly because the debate has been muddled by contourists' adoption of other hypotheses which are not logically consequent on the English contours hypothesis, and by their belief that the levelists' position forced them to deny these apparent truths.

At least four other hypotheses are relevant. These have not always been clearly distinguished from one another.

6.1 The obligatory part of a tonal morpheme is a phonological segment, i.e. a phonological unit which does not contain other phonological units.

6.2 The obligatory part of a tonal morpheme is a phonological unit.

6.3 The obligatory part of a tonal morpheme is a unit.

6.4 A tonal morpheme is a unit.

None of these statements is entailed by the English contours hypothesis, but some contourists would nevertheless maintain all of them, while virtually all contourists would agree with at least one. On the other hand, denial of the English contours hypothesis does not entail the denial of any of these

hypotheses, contrary to the assumptions of some contourists. For the most part I will not attempt to link the various positions with particular contourists. The general lack of explicitness in their accounts makes it futile to do so. My primary aim is to make clear what follows or does not follow from what. I will take the four statements above in turn.

1.1. Phonological segment

As far as I can tell, the truth of the English contours hypothesis follows from the truth of 6.1, given that nuclear syllables are associated with the obligatory parts of a tonal morpheme (which others call a nuclear tone) and that linguistically significant pitch changes occur in the course of nuclear syllables. If nuclear tones are phonological segments, then English uses atomic contours. So 6.1 entails the English contours hypothesis. The English contours hypothesis does not, however, entail 6.1. Vanderslice and Ladefoged (1972) claim that 6.1 is true, and Ladd (1980:189-192) seems to take this position, too. It is difficult to say what proportion of contourists agree, because few of them make their views explicit, but many of them describe the obligatory part of a tonal morpheme as a *tone* or *nuclear tone*, terms which to some linguists imply that it is a phonological segment. I will now show that 6.1 is false, whether the English contours hypothesis is true or not.

At first sight, the introduction of atomic contours provides a way of preserving the hypothesis that nuclear tones are phonological segments, i.e. that they do not consist of sequences of smaller phonological units. Certainly, it would enable us to represent a simple fall as a segment, but it will not work as a general rule for all nuclear tones. Whether we permit atomic contours or not, the obligatory part of a tonal morpheme must be described as a sequence of segments, as I will now show. Suppose, for example, that we represent a falling tune as [+fall, -rise] and a rising one as [-fall, +rise]. Both a rise-fall and a fall-rise would then be [+fall, +rise]. We could distinguish them, following Wang (1967), by means of a feature such as [convex], so that a rise-fall would be [+fall, +rise, +convex], and a fall-rise [+fall, +rise, -convex]. This solution is not simply ad hoc. The feature [convex] has the effect of an instruction to order the two other features in a particular way. It is being used to conceal, rather than

reveal, the fact that [fall] and [rise] are not features of one segment, but features of two subparts of it.

One could avoid this particular problem by not recognizing a rise-fall, but the rise-fall-rise of the Switchback is more difficult to avoid and even more problematical. The feature [convex] does not help here, since the contour seems to be both [+convex] and [-convex]. The simplest feature analysis in terms of atomic contours is [+fall, +rise, +rise]. Again, these are clearly features of ordered subsegments rather than of a single segment. As a characterization of one segment it is completely illegitimate, because a single segment cannot be specified twice for the same feature. It is unlikely that a linguist would make this sort of proposal in such an explicit way, but a linguist could very well do so more subtly, by renaming one of the features, replacing [+rise] with something like [+emphasis]. But if the latter is really an atomic contour feature, labelling it differently is misleading, to say the least. If it is not an atomic contour feature, such an analysis is open to the charge that it treats only some rising contours as atomic contours in an apparently arbitrary manner.

I conclude that nuclear tones should not be described as phonological segments, that is, as single bundles of feature specifications. They must, for the sake of consistency, be analyzed as sequences of segments, whether atomic contour features are permitted or not. One could, of course, preserve the phonological segment hypothesis by introducing a relatively large number of atomic contours, so that a fall-rise, for instance, would be something like [+fallrise, -fall, -rise]. This would be uneconomical, to say the least, in terms of features, and would amount to the claim that no nuclear tone has anything in common, phonologically, with any other one. That this is not what contourists have in mind is clear from the names they use. The name *fall-rise* includes the names *fall* and *rise*. If all nuclear tones were conceived as atomic, moreover, a distinction such as Crystal's between simple and complex tones would not be possible, and such a theory would not distinguish tones which rise at the end from those which do not.

1.2. Phonological unit

Let us now consider hypothesis 6.2, that nuclear tones are phonological

units. This must be true if 6.1 is true, but it need not be false if 6.1 is, as I have argued, false. Hypothesis 6.2 may be true even if 6.1 is false. It seems to be widely assumed by contourists (a) that 6.2 is true, and (b) that 6.2 entails the English contours hypothesis.

It seems fair to say that all linguists who use the term *nuclear tone* take 6.2 to be true. The word *tone* is completely inappropriate for something which is not a phonological unit. As with 6.1, there is a scarcity of arguments offered to support this position. Contourists tend to conclude that it is true on the basis of evidence which might be used to support 6.3, the hypothesis that the obligatory part of a tonal morpheme is a linguistic unit of some kind, but not the more specific claim embodied by 6.2. I have already noted that linguists who maintain 6.2 are usually happy to talk about the meaning of nuclear tones. This is ironic. If it were true that these units have meanings, this fact would strongly suggest that they are not phonological units. The hypothesis that phonological units do not have meaning is very well established.

Even if good support were found for 6.2, this would not affect the debate on the English contours hypothesis. Unlike 6.1, 6.2 does not entail the English contours hypothesis. In other words, it is possible to deny that atomic contours are needed without denying that the obligatory part of a tonal morpheme is a phonological unit. Accepting the levels hypothesis obliges us to claim that some nuclear tones contain more than one phonological unit. This does not mean that the nuclear tone cannot be a phonological unit. Syllables appear to be phonological units, and yet they contain other phonological units, viz. segmental phonemes. Thus students of tone languages can argue for the analysis of so-called contour tones into sequences of level tones without denying that these sequences constitute larger phonological units.

1.3. Nuclear tones as units

Few contourists indeed would disagree with hypothesis 6.3, that the obligatory part of a tonal morpheme is a unit. Certainly all those who call it a tone of some kind would agree. The truth of 6.3 is not, however, self-evident. Many levelists do not recognize a unit comparable to the nuclear tone. Since I have shown in Chapter 1 that there is little support for 6.3, it need not occupy

much of our time here. It will do no harm, though, to make it clear that this distribution of positions is a matter of history rather than logic. On the one hand, as I mentioned above, the English contours hypothesis does not entail 6.3. Recognizing atomic contours does not compel one to recognize the obligatory part of a tonal morpheme as a unit. Thus a contourist is not obliged to insist on the recognition of nuclear tones. On the other hand, 6.3 does not entail the English contours hypothesis. We can recognize a unit like the nuclear tone without abandoning the idea that it consists entirely of level tones.

1.4. Tunes as units

Insofar as these labels imply that a levels theory does not recognize contours as units, they are misleading. Trager and Smith (1951), whose treatment exemplifies the levels theory, make it perfectly clear that pitch levels are constituents of morphemes. They distinguish two kinds of suprasegmental morpheme. "Those consisting of pitches and a terminal juncture", they say, "are called INTONATION PATTERNS" (1951:56).

Trager and Smith therefore propose that the tune associated with a phonological phrase is a unit, but this view does not distinguish levels analysts from contours analysts. There are levels analysts, such as Nida (1949:62), who consider the terminal glide a separate morpheme. There are also some contours analysts, such as O'Connor and Arnold, who recognize a unit corresponding to Trager and Smith's intonation pattern and our tonal morpheme.

The notion of a fairly small inventory of tunes is closely associated with contours theories. There is a widespread assumption among contour analysts that levels analysts deny the existence of such an inventory, but this is not so. The reason Trager and Smith, for example, do not present such a list is a matter of methodology. From their point of view it is unscientific to proceed in this way, i.e. by making hypotheses. It is premature, they would say, to identify morphemes before one has identified the phonological units and then the permissible combinations of phonological units. These combinations are morphs. Only after the morphs were isolated would one proceed to determine

which of these morphs were allomorphs-- variant forms-- of the same morpheme. The number of morphs isolated by Trager and Smith's analysis would be quite large-- at least 1500-- but the number of morphemes involved could easily be on the order of ten or twenty. Thus, for instance, they might determine that the morphs 31, 231, 42, 41, 241, among others (ignoring terminal junctures for clarity's sake), belonged to the same morpheme, which would turn out to be the same entity which contours analysts call a falling tune. These would be allomorphs of a single suprasegmental morpheme just as /z, əz, s/ are allomorphs of the English plural morpheme. It is only a lack of time-- and a commitment to a particular view of linguistics-- that prevents Trager and Smith identifying the suprasegmental morphemes of a dialect of English. Levels theorists have no objection to the notion of an inventory of English suprasegmental morphemes.

2. Arguments against contours

The labels are misleading insofar as they imply that we are to have either level tones or contour tones. The choice is rather between a theory which uses only level tones and one which has contour tones as well as level tones. Most, if not all, contourists include at least one "level nuclear tone" in the inventories they propose for English. The levels theory is therefore simpler than the contours theory in the sense that it requires only one kind of tone. From this point of view the introduction of contour tones represents a complication which needs to be justified.

The idea that English has contour tones receives support from the existence of such phonological units in other languages. Indeed, it is unlikely that contour tones would have been proposed for English if they had not already been used by linguists for the description of languages such as Chinese. Up to about 1970, there was substantial agreement among linguists that contour tones were required for the description of several Asian languages. More recently, however, the notion that contour tones are needed for these or any languages has come under fire. For example, Anderson (1978:160) reviews in some detail the arguments for contour tones which have appeared in the

literature and concludes that "there is really no evidence that points clearly to the necessity of representing contour tones as units in any language". There is, on the other hand, evidence that "contour tones must be decomposed into sequences of levels in some languages".

Anderson explicitly confines his attention to tone languages, but his conclusion is certainly relevant to the question of atomic contours in English. The case for atomic contours in Asian tone languages appears to be stronger than the case for atomic contours in English. Most of the arguments considered by Anderson cannot be used for English, and there is a scarcity of arguments in the literature to replace them. It is unlikely, therefore, that the case for contours in English will succeed where the case for contours in Asian tone languages failed. Furthermore, even if there are independent arguments for contour tones in English, their rarity would make them suspect. If contour tones were to be introduced into linguistic theory for the sake of one language, an extremely convincing case would have to be made.

I noted above that English tunes contain tones which are clearly level in addition to possible contour tones. It appears that it is not coincidental that English does not use only contours. According to Maddieson (1977) there are many tone languages which use only level tones, but none which have contour tones and no level tones. From the point of view of a contours theory, this is strange, if not problematical. If there are atomic contours as well as atomic levels, we should expect to find not only languages with only level tones and languages with both types, but also languages with only atomic contour tones.

I assume that a scientific linguist is in the business of comparing theories. There are two ways in which a theory may be inadequate compared to another. It may fail to make predictions where the other one does. In this case we prefer the latter because it is both more comprehensive and more subject to disconfirmation. On the other hand, one theory may make incorrect or less accurate predictions than another. Contours theory is inferior to ours in both respects.

Contours theories fail to make predictions about the distribution of tunes across the syllables of phrases, and the same is true of traditional levels theory.

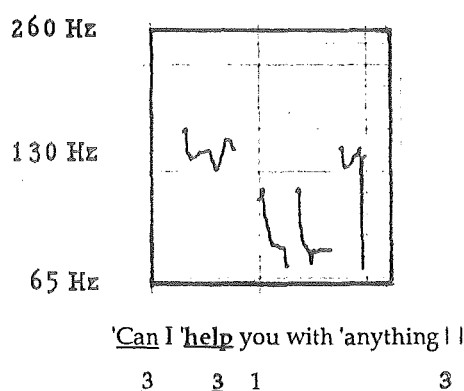
Unlike them, our theory contains explicit tone association rules which attach the tones of tunes to syllables. These rules represent generalizations which hold for an infinite number of phrases. Thus our theory makes a great many predictions where contours theories make none, and the results are certainly encouraging. It is sometimes suggested that tone association rules could be developed for atomic contour tones. It is difficult to see how this could be done without breaking the contours down into constituent level tones, but I look forward to demonstrations, rather than assertions, that it can be done. Autosegmental phonology was invented largely to avoid the problems which arise with the distribution of contour tones across segmental phonemes.

In fact, most versions of contour theory are so vague and informal that they make very few predictions. In order to evaluate them in this way it is necessary to imagine them formalized. The alternative is simply to ignore them. It is not difficult to work out what these theories predict if formalized, but it certainly leaves us open to the charge of distorting the claims of contours theorists. If proponents of such theories wish to demonstrate that we have misrepresented them, they will have to make their theories rather more explicit than they have been.

Contour theories predict incorrectly that a rising tune will be a mirror image of a falling tune. In particular, they predict that the pitch changes involved in a rise or fall are normally distributed equally across the syllables carrying the obligatory part of the tune. We have already seen in Chapter 1 that this claim is incorrect.

Contour theories predict that the change of direction in a fall-rise takes place half-way between the beginning of the nuclear syllable and the end of the phrase. There are actually two predictions here. The first is that there is only one change of direction. In fact, there are often two changes-- from falling to level and from level to rising, for example. When it is possible to speak of one change of direction, it is only rarely at the midpoint. In the following trace, for example, I claim that the syllables between the nuclear and final syllables are phonologically on the same pitch, so that there are two changes of direction. If we consider it phonetically, then there is only one change of direction, but it is

at the beginning of the final syllable, not in the middle of the postnuclear section.



It might be argued that, even if existing contours theories make such predictions, contours theories can be modified so as to make the right ones. The question remains, however, whether they will still be contours theories once these modifications have been made.

Finally, the position of Pike on this question must carry some weight. Since Pike (1948) was one of the first to expound the distinction between level, or register, tone languages and contour tone languages, he has no prejudice against contour tones. Nevertheless, he does not believe that atomic contours are needed to describe English intonation. As we have seen in Chapter 4, he postulates for English four level tones, or pitches.

3. Arguments for atomic contours

Anderson reviews Pike's arguments for atomic contours and notes (1978:154) that the "search for further arguments in favour of Pike's position has not been a particularly fruitful one". The search is even less fruitful with respect to English intonation. One of the reasons for the scarcity of arguments for a contours theory seems to be indifference. To many British linguists the choice of a theory was a minor matter. Crystal (1969a:197) makes this plain when he says that the "problem of choosing a theoretical model is not very great, thanks to a combination of geographical and historical circumstances". That is, he chooses the contours theory because it is British like himself.

Approximation to the truth is apparently not an important consideration.

Since Pike believes that English uses only level tones, his arguments clearly do not apply to English. Anderson reviews three other arguments, which he identifies as (1) the argument from perception, (2) the argument from nonidentity of contours and sequences, and (3) the argument from tone sandhi in Asian tone languages. Only the first of these seems to be applicable to English. This kind of argument goes back at least as far as Sapir, as Anderson makes clear, but Sapir does not apply it to English. Since Anderson's discussion is clear and concise, I will quote him here instead of paraphrasing. Work on tone perception, he says (1978:154),

supports the claim that contours have a high degree of perceptual salience and that human perception of tone is likely to be heavily dependent not simply on tone levels, but on movements. In part, this is simply because the absolute pitch value corresponding to a given tone may be quite variable, and hence the identification of tone level may be quite difficult, while a rise is always easily distinguishable from a fall, regardless of their relative levels.

In considering the weight to be given to this observation in constructing a theory of tone, however, one must be careful to distinguish between the appropriate linguistic description of a phenomenon and the strategies used by speakers and hearers to identify linguistic elements. It is undoubtedly the case that speakers and hearers use contour, where it is available to them, as an aid in identifying tonal elements. It has not been demonstrated, however, that the perceptual salience of contours is correlated with the extent to which a linguistic structure treats them as units.

Anderson goes on to discuss a tone perception experiment reported by Gandour (1978), and concludes (1978:155) that so far

all we have is evidence for a general strategy in perception, applicable to languages for which the linguistic evidence shows clearly that contours are phonologically composed of sequences of level tone units as well as to languages for which such evidence is not available. A strict isomorphism between the categorizations yielded in perceptual experiments . . . and the categories of linguistic structure in a particular language should not be expected to obtain, precisely because languages can clearly differ in many more ways from one another than there is reason to believe humans differ in overall perceptual organization.

As Anderson points out, we find comparable situations in other phonological domains:

It is clear, for example, that vowel length is a major perceptual clue in

English for the voicing of a following stop. No one (or at least not the majority of linguists) would conclude from this fact that vowel length and not stop voicing is the basic phonological difference between, e.g. *muck* and *mug*. If vowel length were the difference, it is hard to see how the distinction between the final segments of inflected forms (*mucked* versus *mugged*, *mucks* versus *mugs*) could be predicted by anything approaching a natural rule.

4. Contourist critiques of levels

Contour theorists typically justify their approach by claiming that the alternative-- levels theory-- is deeply flawed. If we look at the flaws they point to, however, we find that they are either minor blemishes or else flaws in an approach to linguistics rather than in a theory of a particular aspect of English. Let us now consider the arguments presented by Crystal (1969a), Cruttenden (1986), and Ladd (1980).

4.1. Bolinger/Crystal

The criticisms espoused by Crystal are essentially those of Bolinger (1949, 1951), whom he describes as the source of the "most powerful and specific arguments". Crystal summarizes these twice, the second time with explicit approval. For the most part these are directed at the Trager and Smith version. I will simply quote the arguments in the order Crystal (1969: 53-4) gives them in, inserting my comments after each one.

(i) "There are too many differences between segmental phonemes and intonation to allow extrapolation from one analytic method to the other: segmental phonemes are semantically discontinuous, whereas intonation is not; the former are arbitrary, whereas the latter is tied to nervous tension; phonemes have many different articulatory correlates, whereas intonation has only one (fundamental frequency); and so on." In the first place, this is evidently a critique of Trager and Smith's analytic procedures, or the procedures they claimed to use. It is not a critique of a theory of intonation. In short, it is beside the point, and Crystal makes no attempt to explain how it is relevant. Let us nevertheless consider the supporting evidence. The idea that phonemes are semantically discontinuous is difficult to understand, since

phonemes do not have meaning. As for arbitrariness, the actual pitch of a particular tone or pitch is tied to nervous tension, but so is the actual quality, length, degree of aspiration, etc. of a segmental phoneme. As for the third point, Meyer-Eppler (1957) shows that fundamental frequency is not the only acoustic correlate of intonation. But let us grant that tones are different, phonetically and phonologically, from segmental phonemes. It does not follow from this that they are not phonological units.

(ii) Since "hardly any analysis of prosodic data has . . . taken place, . . . the decision to have four pitch phonemes is arbitrary". This seems to be a rejection of the theory because it is underdetermined by the data. In other words, linguists are being condemned for making hypotheses. The number four may be arbitrary, but several linguists seem to have proposed it independently. If others have a 3-tone or 5-tone theory which is better, all they have to do is demonstrate this. Of course, Trager and Smith do talk in terms of facts rather than hypotheses, but this need not confuse us about what they are doing.

(iii) "The notion of 'relative' in relative pitch levels has never been defined, either for individuals or for groups." The problem here is not the notion of intonational morphs consisting of pitches, but the assumption that pitches, like other phonological units, can be identified from phonetic information alone. This assumption is certainly separable from the theory of pitch levels. The hypothesis that the actual pitch of a pitch phoneme can vary so much that the range of separate phonemes may overlap may make the theory difficult to test, but it does not make it wrong. Good scientific theories, whether in linguistics or physics or some other discipline, are usually difficult to test. Crystal does not escape this difficulty. "Granted that none of the pitch levels is absolutely defined," he says, "one would expect the level analysts to make an alternative positive statement about the approximate range of each pitch phoneme" (p. 198). Here the pot seems to be calling the kettle black, since Crystal himself does not provide such a statement about the approximate range of normal, high, and low onset syllables (p. 143-4).

(iv) "There is no scale of linguistic relevance for intonational meanings in levels analysis". The relevance of this is not clear. What is the point of criticizing phonologists and morphologists for not doing semantics?

(v) "A case could be made out for pitch phonemes not constituting morphs". This is true. Crystal does not, however, make such a case. There is always the possibility of a better theory, but this fact does not constitute an argument against an existing theory.

(vi) "Bolinger also supplies a number of examples of utterances which level-analysis either cannot represent, or does so ambiguously." This is likely. I show in Chapter 4 that the Switchback tune on a single syllable is embarrassing for Trager and Smith's system, but it does not follow that no levels theory can represent it.

4.2. Cruttenden

Cruttenden (1986:45-6) says that levels analysis "has been heavily criticised on a number of counts" and briefly discusses four of these. The first is the same as (iii) above: "Proponents of the levels theory have always claimed that it was relative not absolute pitch that was being discussed, but the question is 'How relative?'" The second is the same as (ii) above: "there is no real reason for having four levels rather than, say, three or five".

The third is that "the system does not make clear the tonetic details of the transition from one pitch to another. For instance, in ⁴John didn't do it¹# where exactly does the fall from 4 to 1 take place?" (Cruttenden omits the stress marks which would normally be included in this kind of transcription.) This is true enough, but it is not a sound reason for rejecting such a system, because transcriptions are not expected to show every detail. One might with equal justification condemn a segmental transcription containing /...ana.../ because it does not make clear at what point the velum closes off the nasal cavity. Moreover, the Trager and Smith transcription does not make the tonetic details clear, but neither do prosodic transcription systems using contour tones, like those of O'Connor and Arnold and Crystal. Proponents of contour theories often describe what happens-- often incorrectly-- but so do Trager and Smith (1951:47): "When change is from one pitch level to a lower one, the movement can usually be described as smooth". They are clearly well aware of the inadequacies of their treatment of transition and scope phenomena, but they are not, after all, writing a monograph, or even an article,

on intonation.

The fourth criticism cited by Cruttenden bears on terminal contours and junctures: "Terminals are only significant when they reverse the preceding pitch direction". This is irrelevant as a general criticism of levels theories, because terminal contours are not a necessary part of such theories. Pike (1945), for example, does not use terminal contours in his widely-known levels theory.

Elsewhere Cruttenden (1986:73) also damns Trager and Smith's exposition because it "hardly treats meaning at all". The comments on (iv) above apply again. Unless there is a competing theory which offers a better explanation for similarities and differences in meaning, such a complaint is irrelevant. Cruttenden refers to heavy criticism of levels analysis. When we examine the arguments, though, we find that they are far from weighty, and mostly far from the point.

Cruttenden (1986:47) admits that most of the arguments against levels analysis apply to contours analyses as well, but he opts for a contours approach, nevertheless. The reason which he cites as the deciding factor "which tips the balance towards a contour analysis" is hardly convincing. This is his "conviction that there is some basic similarity of meaning in all falling contours as opposed to a basic similarity of meaning in all rising contours". First of all, this is surely an odd decisive criterion in view of our knowledge of the semantics and pragmatics of intonation. Our understanding of semantics and pragmatics is not firm enough to support such far-reaching decisions about phonological representations.

But let us grant for the sake of argument that Cruttenden is correct about the semantics of falling and rising contours. He claims (1986:47) that an "analysis in terms of levels does not capture this basic division between falls and rises". This is not true. Levels theories can indeed capture this division. This applies both to "mixed" theories like Trager and Smith's, which use terminal contours in addition to pitch levels, and to "pure" theories like mine, which do not use terminal contours. In fact, it is Cruttenden's contour analysis which does not capture this division.

Let us begin with "mixed" levels theories. Cruttenden (1986:47) has this to say:

Proponents of a levels analysis might object that this is precisely what terminal junctures do. But it has been previously noted that it is impossible to isolate such terminals where they merely continue a preceding pitch movement, e.g. in cases like 3-1# (falling) or 2-3 | | (rising). So terminals may capture the basic difference between rising and falling but ascribe the difference to a portion of utterances which is difficult to isolate.

The words "it has been previously noted" are rather deceptive. Cruttenden does not show, either here or previously, that it is impossible to isolate such terminals. On the preceding page he makes a different claim: "Terminals are only significant when they reverse the primary pitch direction". This seems to be quite a different claim, but it matters little, because no justification is provided for it. Neither does Cruttenden explain why he thinks the difficulty of isolating a linguistic item is such a damning argument against it.

Theoretical entities, like phonological units and features, are by their very nature difficult, if not impossible, to isolate. What Cruttenden needs to do is show that levels analysts transcribe some rising tunes without rising junctures or some non-rising tunes with rising junctures, or both. This he does not attempt to do.

Cruttenden seems to assume that a levels analysis is bound to use terminal junctures. This is certainly not true. Moreover, a "pure" levels theory, in which tunes consist only of pitch or tones, is perfectly capable of capturing formally the distinction between falling and rising contours. In the former the final tone is lower than the one before it; in the latter the final tone is higher than the one before it. Cruttenden does not mention the level contour in this discussion, and its omission is understandable, because it is an embarrassment to a contours analysis like his. His theory predicts-- if we force it to predict-- that a level contour will be semantically quite different from both falling and rising contours, since it has nothing in common formally with either of them. In fact, it is widely agreed by contours analysts that level contours belong semantically with rising contours. This is not embarrassing to a levels theory, because such a theory is not committed to a three-way

distinction in terms of form. In a levels theory, a two-way distinction is easier to specify than a ternary one. Falling contours are all those tunes in which the final tone is lower than the penultimate (obligatory) tone. Level and rising contours are all those in which it is not. In my own levels theory, the distinction is even easier to specify, because all falling tunes, and only falling tunes, end with a 1-tone. All those which do not, therefore, form a natural class. (A natural class is a class of entities which can be identified more easily in a particular system than any of its members.) In short, there is a formal common factor amongst level and rising contours to which we can, if required, attribute their alleged functional or semantic resemblance.

Cruttenden claims to reject levels analysis because it cannot capture the basic division between falls and rises. Cruttenden does not make it clear what he means by this, but as far as I can see, he wants a theory in which all and only falling contours-- for example-- have something in common formally, because they have something in common semantically. We have just seen that levels theories can in fact meet this criterion. Cruttenden takes it for granted that a contours analysis does meet this criterion, but does not demonstrate that it does. This is not surprising, because his contour analysis is itself incapable of meeting his decisive requirement.

Cruttenden's version of contour theory is typical in offering no formal analysis of the tunes. Nuclear tones do not consist of smaller units. The name Fall-Rise contains the name Rise, but names are not formal representations. The diacritic \sim in a sense contains the diacritic \acute , but diacritics are not formal representations, either. If we consider the theory, rather than the names and diacritics, each nuclear tone is equally different from all the others. A Fall-Rise shares nothing with a Rise that it does not share with a Fall. There is therefore no common formal factor amongst rising tunes to which one could attribute their semantic resemblance.

Finally, Cruttenden contradicts himself by opting for a levels theory elsewhere in his book. In his concluding remarks he recognizes that what he calls tones do need to be represented as sequences of levels (1986:183-4):

It is likely that the tones will have a dual formal representation: contours, to which the abstract meanings . . . are attached, but also

sequences of levels, which are specified for each contour because this makes possible a simpler statement of the realisation rules . . . involved in mapping the tones from the intonational lexicon onto varying stretches of segments which have pre-assigned stresses . . . and nucleus placements.

4.3. Ladd

In this section I will deal with the anti-levels argument presented in Chapter 8 of Ladd 1980, which is essentially the same as Ladd 1978. It seems to me that virtually all the steps of Ladd's argument are weak. I will attempt to summarize it and then go on to criticize it step by step.

Ladd (1980:185) posits

the existence of a general phenomenon of 'stylized intonation', which is used to signal that an utterance is in some way part of a stereotyped situation or is otherwise more predictable or less informative than a corresponding utterance with plain intonation. Stylized variants are characterized by level pitches: stylized fall is a stepping-down sequence of two level pitches, and stylized rises . . . are a single level pitch.

He then argues that a pitch-level analysis is inadequate because it is unable to express generalizations based on pitch direction, and the phenomenon of stylized intonation requires a theory which can express such generalizations.

4.3.1. Stylized intonation Ladd claims that the calling tune, which I discussed in Chapter 5, is a stylized fall. That is, it is related in a particular way, functionally and formally, to the tune which I represent as (1)(3)3̇1. Ladd supports this claim by arguing that there are other pairs of tunes related in the same way. Thus he claims that (1)(3)2̇-- O'Connor and Arnold's Terrace, apparently-- is a stylized variant of his low-rise, and that the (1)(3)4̇ tune is a stylized high-rise. The following table will make this clear.

	Plain	Stylized
Fall	(1)(3)3̇1 [Drop]	(1)(3)3̇2
Low-Rise	(1)(3)1̇3 [LowBounce]	(1)(3)2̇ [MidLevel]
High-Rise	(1)(3)4̇5 [HighRise]	(1)(3)4̇ [HighLevel]

I will refer to Ladd's plain and stylized low-rises as Low Bounce and Mid Level respectively, and to his plain and stylized high-rises as High Rise and High Level. I have identified the obligatory tone of his stylized low-rise as a 2-tone because the syllables attached to such tones in his transcriptions are not as low as the lowest syllables. As a matter of fact, Ladd expresses his claims in terms of nuclear tones rather than tunes, but they are easily translated, since all of his examples of each nuclear tone exemplify only one of our tonal morphemes. For instance, his low-rise can be identified with O'Connor and Arnold's Low Bounce because none of his examples illustrate the Take Off--they all have at least one strong high syllable before the nuclear syllable. Since Ladd does not show that the Mid Level corresponds to his low-rise irrespective of the nature of the prenuclear syllables, it seems appropriate to talk in terms of whole tunes rather than nuclear tones.

4.3.1.1. Let us begin by examining the claim that the calling tune is a stylized version of the ordinary fall. Ladd does not so much demonstrate that the calling tune is related to the fall as assume it. This assumption seems to be based on the fact that a fall can alternate with the calling tune without changing what he calls the basic function of the utterance, as in his example 20 (1980:177). (I will continue to omit tone association lines where no syllable is attached to more than one tone, placing each tone under the leftmost syllable to which it is attached.)

A: [from a distance, pointing to the car from which B has just emerged]

You 'left your 'lights 'on

1 3 3 2

B: [who was jangling keys getting ready to lock the door] What?

A: [louder] You 'left your 'lights 'on

1 3 3 1

This might be significant if the calling tune were the only tune which could be

replaced by a fall, but the fact is that a fall can replace virtually any tune without changing the basic function. For example, it can replace a High Rise. We can construct another exchange in which a sentence with a High Rise is repeated with a Fall. It does not follow from this that there is a special relationship between these two tunes (or nuclear tones).

A: [who had seen B approach on foot, then apparently return to his car, before finally arriving at A's house]

Did you 'leave your 'lights 'on?

1 3 4 5

B: [who is now thinking about something else] What?

A: Did you 'leave your 'lights 'on?

1 3 3 1

Or again, Ladd points out that the calling tune is appropriate for warnings that are essentially reminders, and the falling tune for warnings in emergencies. It does not follow from this that the two tunes are closely related. The Low Bounce can also be used for warnings that are essentially reminders, but one would not want to argue that the Low Bounce is therefore a stylized version of the Fall.

Look 'out for the 'broken 'step.

1 3 1 3

So the Fall alternates with tunes other than the calling tune without changing the basic function of the utterance. Ladd's assumption can also be attacked from the opposite direction. That is, the calling tune alternates with tunes other than the Fall. The most obvious candidate is the Dipper, which, like the calling tune, is frequently used with *thank you*, *thanks*, and *excuse me*. And, as Cruttenden (1986:125) says, both the calling tune and the Dipper (Fall-Rise nuclear tone in his terms) "are commonly used as 'call contours'".

In fact, we can perhaps gauge the weakness of Ladd's argument by the fact that both Cruttenden and Bolinger (1986:226-234), who accept the notion of stylization (or monotone, in Bolinger's terms), describe the calling tune as a stylized version not of the Fall but of the Fall-Rise, i.e. Bolinger's Profile AC. Indeed, Cruttenden appears to attribute this view to Ladd.

Another problem for Ladd's hypothesis is the newspaper vendor/conductor chant mentioned in Chapter 5 above. If the calling tune is a stylized fall, then so is the conductor tune. If the plain fall has two stylized versions, we should expect other tunes to have two stylized variants as well, but this does not appear to be true. Ladd therefore attempts to sidestep this incorrect prediction by claiming that these two tunes differ on a gradient dimension. We have seen, however, that the distribution of the constituent tones over syllables is clearly different, with the final tone of the conductor tune beginning on the first postnuclear syllable, and that of the calling tune beginning on the first strong postnuclear syllable. I would have thought that this was an example of a formal difference which was clearly not gradient-- a yes or no matter, not a more or less one.

4.3.1.2. Ladd's stylized intonation argument also depends for support on his claim that the High Level is a stylized version of the High Rise. This in turn depends heavily on the following claim (1980:183):

Plain high rise is very commonly used for listing in English:

(40) milk egg^s butter b^read
 I need and and and

Ladd goes on to contrast this with the High Level (pp. 183-4).

Stylized high-rise-- with the rise becoming steady level-- is also common in lists, with the implication 'etcetera'. That is, the items in the list are not individually informative, but rather are intended to suggest a loose grouping which the hearer can fill out for himself. Thus:

(41) A: Hey, these cookies are good. What's in 'em?

B: Oh, nothing special, you know--

flour-- sugar-- butter--
and and and, uh ...

This 'etcetera' use of stylized high-rise in lists shows up frequently in casual conversation:

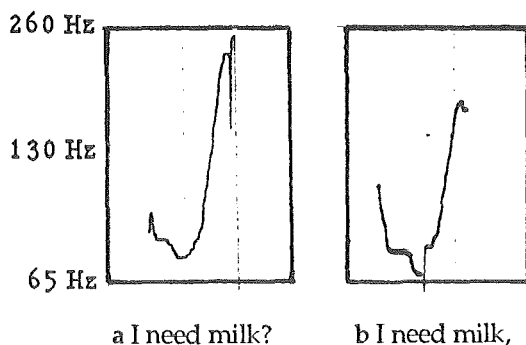
(42) A: What a ridiculous day I've had. I spent the whole morning running around downtown.

B: What for?

A: had to get my car registered-- stamps--
I and buy some and ...

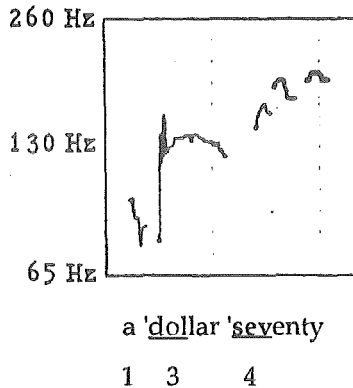
It doesn't matter exactly what A had to do; the general idea is 'time-consuming errands'.

The claim that plain high-rise is very commonly used for listing is highly questionable. I think that Ladd's example 40, like 41 and 42, illustrates the High Level. If a High Rise is used on *I need milk*, it sounds like an echo question-- i.e. *Did you say that I need milk?* -- rather than the beginning of a list. Compare the following traces. In the first, *I need milk* carries a High Rise; in the second it carries a High Level.



These are clearly not homophonous, since the final pitch is considerably higher in a than b, but phonetically speaking, the pitch is rising throughout both instances of *milk*. Presumably this is what leads Ladd to claim that b exemplifies a plain high-rise, i.e. a High Rise rather than a High Level. But b is a perfectly normal manifestation of the High Level tonal morpheme on a phrase with no postnuclear syllables. In Chapter 4 this morpheme was

illustrated with the phrase *a 'dollar 'seventy*, where the nuclear syllable contains, like *milk*, a short vowel. I reproduce the trace here.



Phonologically, the last three syllables are attached to a 4-tone, but the phonetic pitch level associated with this tone in this utterance is not reached until the end of the nuclear syllable. It does not level out until after the nuclear syllable. It is not surprising, therefore, that phonetically, the pitch of *milk* in Ladd's example is rising, especially since it has to get not from tone 3 to tone 4 but from tone 1 to tone 4 within the space of a very short syllable. (I assume that the prenuclear syllables in Ladd's example are low, since he represents them on the same level as the following *ands*, which would normally be weak and low, but this is of little consequence.) A relatively high-pitched instance of the High Level could certainly be homophonous with a High Rise, but the longer the part of the phrase attached to the nucleus tone is, the less likely this becomes. It is interesting in the light of this to note that all three of the relevant nuclear syllables in Ladd's 40-- *milk*, *eggs*, and *but*-- contain phonologically short vowels, and two of them end with vowel-shortening voiceless consonants. Moreover, only one of these syllables, *but*, is followed by a postnuclear syllable, and it is weak and therefore short, as well.

I conclude, then, that Ladd's example 40 illustrates not a High Rise, but a High Level. The latter does not, therefore, alternate with the High Rise in lists, and is not a stylized variant of it. This is not to say that the relevant parts of Ladd's 41 and 42 must be prosodically indistinguishable from 40, but only that they do not manifest different tonal morphemes. The differences may be attributed to lengthening of the nuclear syllables, or perhaps entire words, for

specific purposes. Such lengthening is certainly not constrained by the tonal morpheme attached to a phrase. In Ladd's 41, the words are being lengthened to buy time to remember what is in those cookies. This could be produced by a stretching rule, along the lines suggested by Bolinger (1972a) or Kuiper (1984), or by a change in tempo. In 42, the extra length appears to be iconic. That is, taking extra time to say the words *car registered* and *stamps* is an effective way of emphasizing that these errands were, as Ladd himself says, "time-consuming".

Ladd also claims that the High Level alternates with the High Rise in questions, where the High Level has an "overtone of tiredness or resignation".

can I 'go 'now

1 4 5

can I 'go 'now

1 4

This is unconvincing because the same question can be asked with a Fall. One could therefore argue by the same token that the High Level is a stylized Fall.

4.3.2. Pitch direction generalizations According to Ladd (1980:186), the phenomenon of stylized intonation is particularly significant because it "sheds new light on the old levels vs. configurations debate". I have shown that the case for stylized intonation, on which Ladd's anti-levels argument depends, is not very strong, but even if we grant the stylized tone argument, the anti-levels argument does not hold up.

Ladd (1980: 190) claims that a pitch-level analysis is "in principle unable to express generalizations based on pitch direction", and that the "relationship of stylized and plain intonation involves generalizations based on pitch direction". Apparently these generalizations are "that falling contours are stylized as a sequence of two relatively steady level pitches, while rising contours (either high or low) are stylized as a single level pitch". In the first place, Ladd does not mention more than one kind of falling contour, and certainly does not show that all falling contours are stylized in this way. In fact, it is not clear whether he recognizes more than one kind of falling contour. The first generalization therefore applies to only one item. Surely any theory

whatsoever is capable of expressing a generalization which applies to only one linguistic entity. Turning to the second generalization, we note that Ladd's theory can pick out the two rising contours because they share a distinctive feature [Rise]. He does not say whether the fall-rise contour (for example) shares this feature. If it does, he has no explanation for the absence of a stylized fall-rise. If it does not, he has no explanation for the fact that the rise is manifested in the same way-- on the final syllable. This is, after all, supposed to be a phonological feature.

It is not clear why Ladd says generalizations based on pitch direction are beyond a pitch-levels theory. Distinguishing rising contours in such a theory is perfectly simple. If there were, as Ladd would have it, no "plain" level nuclear tones, all basic tonal morphemes would have more than one obligatory tone. Rising tunes would then be all and only those tonal morphemes in which the final tone was higher than the penultimate one. This would include so-called complex rises such as the Dipper and Switchback. We could exclude these, if required, by referring to all and only those tonal morphemes with only two obligatory tones in which the second of these is higher than the first.

Furthermore, Ladd's claim that we require a generalization in terms of pitch direction is open to question. He assumes that the low-rise and the high-rise are stylized differently from the fall because they are rising contours. There are certainly other possibilities. For example, it may be because they both, unlike the Plain Fall, contain end tones. In fact, it can be argued that the latter explanation is better. The argument goes as follows.

In a contours theory like Ladd's, the only difference between rising and falling tunes seems to be the difference in direction. There is nothing in the representation of these tunes to account for the fact that rising tunes are stylized with one tone, and the falling tune with two. In our levels theory, however, the rising tunes end with an end tone, and the falling one does not. If we suppose that the stylized tunes are derived from the corresponding plain ones by a rule which deletes end tones and replaces 1-tones with 2-tones, the difference between stylized falls and stylized rises would follow in a natural way from an independently motivated difference between plain falls and plain rises.

	Plain	Stylized
Fall	(1)(3) <u>3</u> 1	(1)(3) <u>3</u> 2
Low-Rise	(1)(3) <u>1</u> 3	(1)(3) <u>2</u>
High-Rise	(1)(3) <u>4</u> 5	(1)(3) <u>4</u>

It seems, therefore, that one aspect of the phenomenon of stylized intonation is handled better by a levels theory than by a contours theory. There is another aspect of stylized intonation which makes trouble for a contours analysis like Ladd's. If his hypothesis is correct, not all tunes have stylized variants. Apparently, only three of them do. It is not clear whether a contours theory can pick out these three from the other plain tunes in a straightforward way. In my levels theory, though, these are all and only those tonal morphemes which have two obligatory tones. (Actually, this would include the Take Off, which is either (1)13 or (1)(1)13. We can exclude it, if necessary, by specifying all and only those tonal morphemes with an accent 3-tone which have two obligatory tones. Of course, excluding the Take Off while including the Low Bounce would be more difficult for a contours theory, and impossible for a theory in which stylization is a property of nuclear tones, since the obligatory elements of the Take Off and Low Bounce morphemes are the same.)

Ladd (1980:191) is critical of levels analyses because there is nothing in the way that (1)(3)31 is related to (1)(3)32 to suggest that (1)(3)13 and (1)(3)2 or (1)(3)45 and (1)(3)4 exhibit the same relationship. The explicitness of such theories makes their shortcomings clear. Ladd's position is more difficult to criticize, because he does not present any formal representations of the tunes or nuclear tones in question. He proposes to capture the relationships in a way which is ad hoc: the stylized tunes are [+Stylized], and the plain ones are [-Stylized]. It is far from clear, however, that this feature, introduced to account for alleged semantic correspondences, is a phonological feature. In the second and third parts of his chapter, Ladd presents data which could support the postulation of a semantic or pragmatic feature [Stylized], but he does not demonstrate the need for a parallel phonological feature. Neither does he

address the awkward fact that other kinds of phonological units do not appear to have stylized variants.

In conclusion, it seems fair to say that the phenomenon of stylized intonation, if it exists, does not demonstrate the inadequacy of a levels approach to English intonation, but rather shows its superiority over an approach which characterizes contours as atomic.

I have in this chapter defended the claim that English tonal morphemes consist entirely of level tones, attacking the hypothesis that they contain atomic contours. I began by disentangling the atomic contour hypothesis from several others which are traditionally tied up with it. I then summarized the arguments against contour tones, and those in their favour, showing that the latter are both scarce and unconvincing. Finally, I reviewed in some detail three critiques of the levels hypothesis, in order to demonstrate that they are not particularly cogent.

Chapter 7

Binuclear Tunes

It has been assumed thus far that every English tonal morpheme contains one and only one nucleus tone. I shall call this assumption, whose generality makes it attractive to linguists, the mononuclear tune hypothesis. Some linguists have proposed to complicate the theory of English intonation by abandoning this hypothesis and postulating tonal morphemes which contain more than one nucleus tone. In particular they have claimed that it is necessary to have binuclear tunes-- tonal morphemes with two nucleus tones. In this chapter I will show that this complication is both unnecessary and unable to account for some facts which are explained in a natural way by a simpler theory which maintains the mononuclear hypothesis.

It is worth noting that even linguists whose models include compound tone units recognize the attractiveness of the mononuclear hypothesis. It is so attractive that it leads O'Connor and Arnold, for example, to contradict their own presentation. "By definition", they say (1973:33), "the nuclear tone occurs on the last important word in the word group", a word group being "a group of words which is coterminous with and modified by an intonation tune" (287). The definite article before "nuclear tone" makes it clear that each word group is assumed to carry only one such tone. Crystal, too, states that "Every tone unit contains one and only one nucleus" (209), though he sidesteps direct self-contradiction by equivocating with the term "nucleus". We find a similar tension in more general works, too. In *Modern Phonology* (1977:37), Sommerstein says that "in each intonation contour one syllable, the *nucleus* or *centre*, is stressed much more heavily than the remaining strong syllables". He then proceeds to contradict this by claiming that some intonation contours contain more than one such syllable:

a sentence . . . may contain more nuclear stresses than intonation contours, e.g. (19) contains (on the most likely rendering) two intonation contours but three nuclear stresses:

19. I *sáid* that *cáts éat báts*, *nót* the *óther wáy róund*.

The mononuclear hypothesis is worth maintaining not just because of its attractive generality, but because abandoning it leads to undesirable complications elsewhere in the theory, viz. in the accentuation rules. Because the proponents of binuclear tunes assume that the number of tunes is always the same as the number of phonological phrases, they must also postulate the existence of binuclear phrases-- phonological phrases which contain two nuclear syllables. That is, they must abandon the mononuclear phrase hypothesis, and this creates serious difficulties. The accentuation rule which makes syllables nuclear refers to the last focussed word in a phrase. It depends crucially on the location of phrase boundaries, and will certainly not provide more than one nuclear syllable per phrase. None of the proposals for binuclear tunes and phrases reviewed in the following pages offers any solution to this problem, or even any discussion of it. It is not necessarily true, of course, that there is a one-to-one relationship between phrases and tunes, but I will not question this assumption in this chapter.

Not all British linguists abandon the mononuclear hypotheses. Brazil, Coulthard and Johns (1980), for instance, appear to maintain them. Neither the supporters nor the opponents of binuclear tunes, however, make any attempt to show why their theory is better than the other. Indeed, Brazil, Coulthard and Johns's reference to "disturbing . . . disagreements over the phonological facts" (8) suggests that they do not think they are doing theoretical work. Phonological facts are not data. Disagreements over physical data might be disturbing, but divergent claims about phonological facts are normal and necessary for the development of better theories, i.e. closer approximations to the facts of linguistic knowledge.

In the first of the following sections I summarize three proposals for binuclear tonal morphemes-- those of O'Connor and Arnold (1973), Halliday (1967), and Crystal (1969). In section 2 I consider three possible reasons for complicating the theory of intonation with binuclear tunes, and demonstrate their inadequacy. In the third and final section I briefly present some positive

reasons for maintaining the mononuclear tune hypothesis, with particular reference to fall-plus-rise utterances, which represent the only kind of binuclear tune recognized by all of the analysts mentioned above. I show first that the tune or tunes attached to such utterances contain two nucleus tones and nuclear syllables, and then that there is a phrase boundary between these syllables, as predicted by a representation consistent with both mononuclear hypotheses.

1. Three proposals

In this section I will outline first the proposal of O'Connor and Arnold, who discuss only one kind of binuclear tune. Then I will deal briefly with Halliday, who proposes two kinds, before taking up at some length the position of Crystal, who allows several kinds of "compound tone". In the course of these discussions I will show how the utterances in question might be represented by a mononuclear theory, and justify these representations.

1.1. O'Connor and Arnold

O'Connor and Arnold (1973:28) discuss only one kind of *compound tune*, which they define as a "tune which contains two nuclear tones". The High Dive, as they call their compound tune, "consists basically of a High Fall followed by a Low Rise". These are the names of nuclear tones. We can represent the High Fall as 31, and the Low Rise as 13. This tune gets attached to one phonological phrase, or word group in their terms, but it contains two nucleus tones: (1)(3)3113.

According to O'Connor and Arnold this tune "may be very similar to some forms of the simple tune containing the Fall-Rise nuclear tone" (p. 28), i.e. the Switchback: (1)(3)1313. It may also be homophonous with the Dipper tune, (1)(3)313, but O'Connor and Arnold ignore this, for the good reason that this tune is not discussed in their book, which is essentially a pedagogical work. As they point out, "*in this book* [italics DH] the Fall-Rise always has the falling head before it" (p. 30).

O'Connor and Arnold make it reasonably clear why the High Dive must

be distinguished from the Switchback, illustrating this with two different intonations of the same sentence, *My mother was born in Sheffield*.

My [˘]mother was born in Sheffield | | (but not my father.) *Switchback*

My [˘]mother was ɔborn in ɔSheffield | | (Isn't that interesting.) *High Dive*

They do not, however, attempt to explain why the High Dive must be distinguished from a sequence of their High Drop and Take Off tunes, which would have exactly the same pitch pattern.

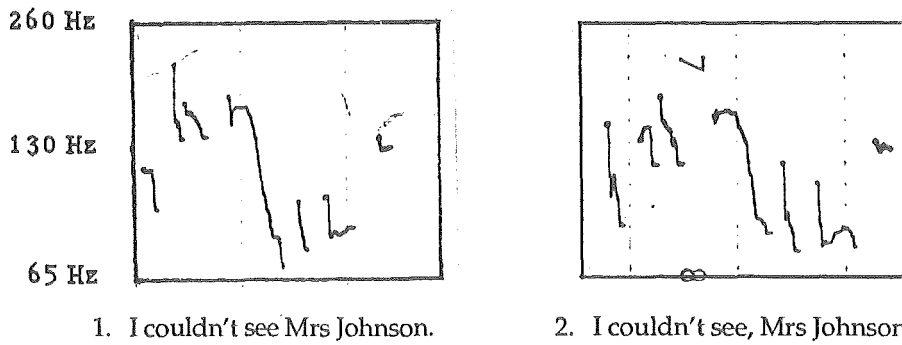
My [˘]mother | | was ɔborn in ɔSheffield | | *High Drop + Take Off*

In fact, O'Connor and Arnold provide a large number of examples of both High Dive utterances (pp. 232-241), and High Drop/ Take Off sequences (pp. 268-269). None of the sentences transcribed with the High Dive contains a comma. On the other hand, the great majority of sentences transcribed as two tone groups do contain a comma, and it coincides with the indicated phonological phrase boundary. Where the indicated boundary does not coincide with a comma, it immediately precedes a sentence-final adverbial such as *tomorrow, sometime, in here, by then, occasionally*, and so on. Moreover, the symbol which O'Connor and Arnold use to mark these boundaries is a single bar, which "indicates the end of a word group . . . after which there is little or no pause" (p. 289). In other words, they are not claiming there is a pause at these points.

These facts can be explained if we suppose that all of the examples exhibit a sequence of two phonological phrases, and that O'Connor and Arnold's transcriptions are determined by the grammatical structure of the associated words. Where there is a major grammatical boundary of the kind typically marked by a comma in written English, they indicate a word group boundary. Where there is not such a grammatical boundary, they do not.

It is clear that O'Connor and Arnold would transcribe 1, where *Mrs Johnson* is the object, as one phonological phrase, and 2, where it is a vocative, as two. The first would exemplify the High Dive tone group, and the second a sequence of High Drop and Take Off.

1. I couldn't see Mrs Johnson.
2. I couldn't see, Mrs Johnson.



There is no evidence for the claim that these utterances differ in their intonation. The traces above show no significant difference. The syntactic structure of 2 might make a pause before *Mrs Johnson* more likely or more appropriate in an utterance of that sentence, but the presence of a pause does not force this interpretation. An utterance with a pause at that point could still be interpreted as an utterance of sentence 1. Whether we conclude that 2 contains one phrase or two, the phrasing of 1 is the same.

Brazil et al. (1980:7) are right, I think, to claim that compound tone groups are proposed to cope with the problem "that it is at times difficult and sometimes impossible to decide where one tone unit ends and the next begins". Introducing binuclear tonal morphemes obviates the need to make such decisions when there is a shortage of clear phonetic and syntactic criteria. This response allows the transcriber to avoid not only the problem of locating the phrase boundary, but also the question of whether there is a boundary or not. Complications are introduced into the theory, it seems, in order to make the task of transcription easier. It has not been demonstrated that they are necessary to account for the intuitions of untrained native speakers of the language. Finding out what these intuitions are is also difficult, but at least it is a linguist's problem rather than a transcriber's.

1.2. Halliday

Halliday (1967) identifies two types of binuclear tune, which he calls double tonic tone groups and labels 13 and 53. The first of these, in which Tone 1 (falling) is followed by Tone 3 (low rising), is the equivalent of the High Dive. In the second type, the same Tone 3 follows Tone 5, which is rising-falling. Thus the distinction between 13 and 53 follows from the distinction between a

fall and a rise-fall, which O'Connor and Arnold do not make. Halliday's Tone 5 appears to be different from the "intonation pattern indicating condescension or mockery" discussed by Leben (1976:101). Tone 5 has what Halliday calls a high pretonic, whereas Leben's tune has a low one. In a framework like mine, Halliday's Tone 5 could be represented either as a distinct tonal morpheme or as a manifestation of the same tonal morpheme as his Tone 3, in which the pitch of the nucleus tone is raised and the nuclear syllable is stretched. Neither hypothesis is problematical, but the second one is attractive because it correctly predicts that the fall and rise-fall will be considered "the same" by the rules which prevent certain tunes following them in certain syntactic structures. Regardless of how we represent Tone 5, utterances with Halliday's Tone 53, like most of O'Connor and Arnold's High Dive examples, can readily be represented as two phonological phrases, each associated with a mononuclear tonal morpheme.

1.3. Crystal

Crystal (1969a) proposes more binuclear tunes than either O'Connor and Arnold or Halliday. In his system two nuclear tones can occur in a single phonological phrase, or tone unit, provided they are not both falling or rising. Thus he recognizes five types (at least) of compound tone, as he calls such a sequence: $\dot{\text{~}} + \dot{\text{'}}$, $\dot{\text{'}} + \dot{\text{~}}$, $\hat{\text{~}} + \dot{\text{'}}$, $\dot{\text{'}} + \hat{\text{~}}$, and $\dot{\text{~}} + -$. I will restrict this discussion to the first two, since they are the only ones which are at all common. None of the others appears in more than 1% of the phrases in Crystal's data, and some are not even illustrated in Crystal 1969a.

1.3.1. Fall-plus-rise I have suggested that, apart from those which illustrate the contradiction contour, O'Connor and Arnold's High Dives could be sequences of High Drop and Take Off tunes. With Crystal's fall-plus-rise ($\dot{\text{~}} + \dot{\text{'}}$) the situation is more complex. Some are sequences of two phrases carrying the above-mentioned tunes, but many seem to be a single phrase with either the Dipper or the Switchback, i.e. a tonal morpheme which contains the obligatory tone sequence 313. O'Connor (1970:15) appears to agree:

I believe that most universities ... raises the immediate question of

differentiating fall plus rise from fall-rise: I think it very likely that I should want to interpret this as a fall-rise on *most*.

It is not difficult to explain why Crystal transcribes such phrases with compound nuclear tones. His practice follows from his definition of the term *complex nuclear tone* (p. 217):

Here I include all nuclei where there is a change of direction of the pitch movement of a kinetic tone *within a syllable* [italics DH], and only one maximum of prominence. The main categories are the fall-rise and the rise-fall.

We have seen in Chapter 1 that the final rise in pitch of the Dipper (and the Switchback) comes on the last syllable of the phonological phrase. The change in direction from falling to rising takes place within a syllable only when the nuclear syllable is phrase-final. If there are syllables between the nuclear syllable and the last syllable, then it is evident that the rise does not begin in the nuclear syllable. Thus Crystal does not indicate a fall-rise nucleus on such syllables simply because they do not meet his first criterion for complex nuclei: a change of direction within a syllable. When the nuclear syllable is in the last word of a phrase, Crystal is able to convince himself, apparently not having access to physical data, that the rise begins in the nuclear syllable, even when this is not phrase-final. Evidently this is because it would be difficult to support a claim that there is a second "maximum of prominence" in the same word. In other words, Crystal assumes that because his second criterion for complex tones is met-- "only one maximum of prominence"-- the second one is met as well.

This explains why there is a scarcity of nonfinal fall-rises in Crystal's transcriptions. Virtually all of the instances of the fall-rise in Crystal 1969a are in the last word of a phonological phrase. In the passage on page 179, for example, there are four examples of fall-rise nuclear syllables, and every one is in the last word of its phrase. Clearly this speaker uses this tune, whether it is the Dipper or the Switchback, rather a lot. It is not unlikely that he uses it six times here. That is, where the nuclear syllable is in the last word, Crystal represents the nucleus as complex; where the nuclear syllable is not in the last word, he transcribes it as the first element of a compound nucleus, the fall-plus-rise.

<i>Complex</i>	<i>Compound</i>
to	I believe that most universities
the money that's required	on a more modest scale
in the meantime	
the first beginnings	

If our hypothesis is correct, we should expect to find that in Crystal's transcriptions the percentage of fall-rise nuclear tones which occur in phrase-final words is higher than the percentage of, for instance, fall nuclear tones which occur in phrase-final words. Unfortunately, there is not enough transcribed data in Crystal 1969a to determine whether this is true.

Crystal admits "the existence of formal overlapping between complex and compound tones" (p. 220). That is, he recognizes that the fall-rise and the fall-plus-rise are sometimes homophonous, but he explicitly denies the possibility that his fall-plus-rise is "a distributional variant of \sim " (p. 220). His arguments, however, are beside the point here. In the first place, neither O'Connor nor I are denying that the fall-plus-rise is to be distinguished from the fall-rise. Rather, we are claiming that some of the tunes Crystal has transcribed as fall-plus-rises are not. Secondly, he says "there are a large number of examples displaying clear semantic contrasts", but the three examples he cites are either irrelevant or unclear. In each of the following pairs, a fall-plus-rise phrase is contrasted with one in which the single nuclear syllable corresponds to the second nuclear syllable in the binuclear phrase.

I | *thòught* it would *ráin* | | *v.* I | *thought* it would *răin* | |

the | *màn* said he'd \uparrow *cóme* | | *v.* the | *man* said he'd \uparrow *cǒme* | |

These pairs are irrelevant because they are not potentially homophonous.

What Crystal needs to demonstrate is a semantic or formal contrast between a fall-plus-rise phrase and one in which the fall-rise nuclear syllable corresponds to the first nuclear syllable in the binuclear phrase. The other pair he cites is relevant in this sense, but far from clear.

| "yǒu don't know | | (Well, who does, then!)

| "yòu don't knów | | (so why are you saying you do!)

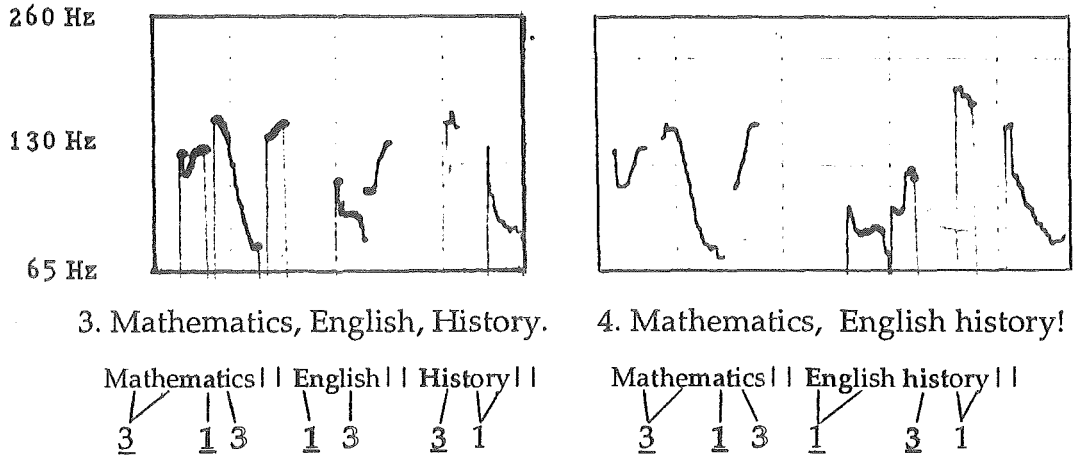
On the one hand, the alleged fall-rise phrase could well be carrying the

Switchback tune, which would not be homophonous with the High Dive, or fall-plus-rise. On the other hand, the second phrase probably exemplifies not the fall-plus-rise but, as I argue in Chapter 8, the tune known as the contradiction contour. Crystal is obviously referring to an utterance used to object strongly.

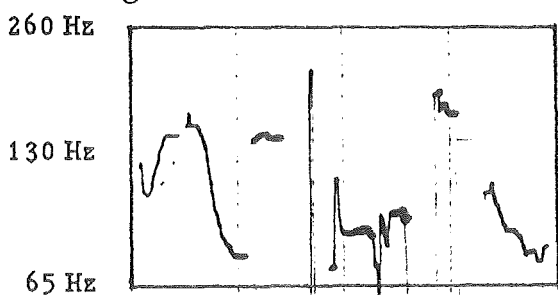
I pointed out above that the transcription problems under discussion stem from Crystal's definition of the complex tone. When we consider his criteria for categorizing nuclear tones as simple, complex, or compound, it becomes clear that most examples of the Dipper and the Switchback-- all those in which the nuclear syllable is nonfinal-- escape his classification. The obligatory portions of the tunes carried by such phrases are not simple because there is a change in pitch direction; they are not compound because there is only one nuclear syllable; and they are not complex because the pitch movement does not change direction within the nuclear syllable. This hole in Crystal's classification system would severely limit the usefulness of the transcription system based on it, if the latter were applied accurately. We are led to the conclusion that when Crystal transcribes High Dive utterances with a fall-rise his transcription is phonetically incorrect in terms of his own system.

1.3.2. Rise-plus-fall Before turning to Crystal's reasons for postulating compound nuclear tones, let us consider the second most common compound nucleus in his data, the rise-plus-fall. As with the fall-plus-rise, I would probably represent some of these as two phrases and some as one phrase. If the utterance consisted of two phrases, the tonal morpheme associated with the first would be either (1)(3)13 (O'Connor and Arnold's Low Bounce) or (1)(1)13 (the Take Off). The tonal morpheme associated with the second phrase would be the familiar (1)(3)31. Most of Crystal's examples, however, appear to be single phonological phrases with only one nuclear syllable. The tonal morpheme involved is (3)(1)31-- the Long Jump. This is to be identified with the "emphatic" version of O'Connor and Arnold's Long Jump.

A phrase carrying the Long Jump tune is potentially homophonous with a sequence of the phrases mentioned above, if the second phrase begins with a nuclear syllable. Consider the following utterances.



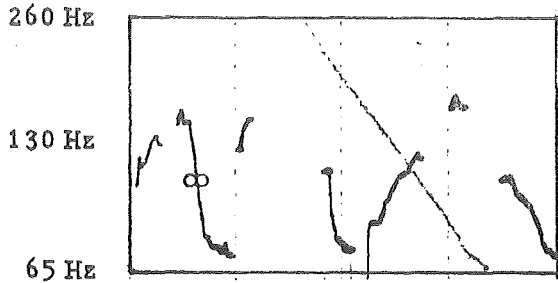
In both of them the syllable before *hist*, the last nuclear syllable, is higher than the preceding syllable. In the first utterance this rise is attributed to the end 3-tone attached to the last syllable of the penultimate phrase. In the second utterance, however, the higher pitch of *glish* is attributed to the 3-tone associated with the following nuclear syllable. In other words, this is an example of anticipatory assimilation. I discuss in Chapter 3 a similar phenomenon in the prenuclear syllables of phrases associated with the Switchback tune. The Long Jump is not always homophonous in terms of pitch with a Low Bounce/Drop sequence, because the prenuclear rise in the Long Jump, unlike that at the end of the Low Bounce, is optional, as the following trace shows.



5. Mathematics, English history!

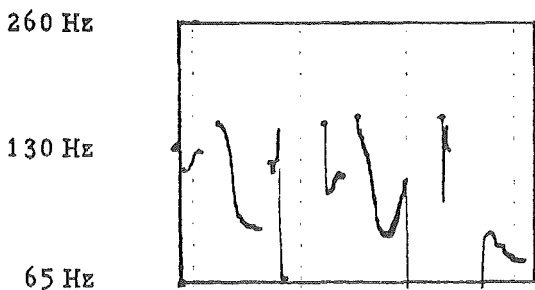
It appears that Crystal would transcribe the second phrase of 4 as a rise-plus-fall, and therefore claim that the first syllable of *English* is nuclear, as it is in 3. In representing the intonation of such utterances as the tonal morpheme called the Long Jump, we make a contradictory claim: the first syllable of *English* in 4 is accented but not nuclear. These claims make different

predictions about what will happen when *English* is replaced with *Japanese*. Since the first syllable of *Japanese* is the first focussed stressed syllable, the accentuation rules will make it accented. Our hypothesis predicts that the first syllable will be more prominent than the third, which is not accented, and that it will be the first low syllable. This prediction seems to be correct.



6. Mathematics, Japanese history.

According to the compound hypothesis, *Japanese* will contain a nuclear syllable, and this must be the third syllable, because it is the primary stressed syllable. This amounts to a prediction that the third syllable will be more prominent than the first. More testably, it predicts that if there is an accented 3-tone-- a relatively high head, as is the norm in Crystal's transcriptions-- then the first low syllable will be *nese*. This is clearly wrong. An utterance like 7 is certainly possible, but (unlike 6) it is intonationally "different" from 4.



7. Mathematics, Japanese, History.

Native speakers would, I predict, hear a phrase boundary after *Japanese*. That is, they would identify it as an appropriate way of speaking a sequence of noun phrases rather than a single noun phrase. A substitution test of the kind described provides an effective-- though not necessarily practical-- means of determining whether a native speaker hears such ambiguous stretches of

speech as one phrase or two-- or sometimes one way and sometimes the other.

Crystal himself provides further evidence for my claim that most of his alleged rise-plus-falls are mononuclear phrases in which only the fall is, in his terms, nuclear. He points out (1969a:219) that extra stress is more likely to occur on the second element of a ´ + ` than on the second element of a ` + ´. This is just what we should expect if only the second element of the ´ + ` corresponds to a nuclear syllable, and the second element of the ` + ´ often does not, as I have already suggested is the case. Crystal explicitly contradicts the suggestion that his rise-plus-fall phrases are mononuclear phrases with the Long Jump tune (p. 223):

Similarly, the ´ + ` cannot be taken as a 'rising head' plus fall, or some other formulation, because of the phonetic characteristics associated with the former:

|áre you gòing|| is very different from |are`you gòing||.

In the former, 'are' is more prominent than (or minimally, equally prominent as) 'going', whereas in the latter it is always less; moreover, the duration of the first tone-unit is much longer than that of the second.

Crystal writes as if he is providing evidence to support his claim, but he is not. He is simply telling us the criteria he thinks he uses to decide whether to transcribe a particular utterance in the first way or the second. We are in no position to contradict this kind of claim, but we might question some of the assumptions behind it. It is far from clear, for instance, that "prominence" is a "phonetic characteristic". Certainly, one could define it as a phonetic term, but Crystal does not do so. Furthermore, he does not explain why the longer of the two types of utterance he cites cannot consist of two tone units. He points out that *are* in the first type can be equally as prominent as *going*. To transcribe such an utterance as a single phrase rather than two is not in line with Crystal's own requirement that "one element of the compound tone must be more prominent than the other" (p. 218).

Crystal's insistence on exploiting compound nuclei leads him to transcribe the word *Japanese* with two nuclear syllables (p. 259):

8. and I |see`by thém|| that |[évery] 'fourth 'child bórn|| `is`a
|jápa↑nèse||

If this is an empirical claim, it is highly questionable. It is very difficult to believe that native speakers would judge the first syllable of *Japanese* in this utterance to be at least as prominent as the third. More significantly, there is no reason for the first syllable to be nuclear here. Our accentuation rules do not predict it. They will make *Jap* nuclear only when it is focussed and the final syllable is not. This would be appropriate if *Japanese* was being contrasted with *Chinese*. In such a context the first syllable might be nuclear. In the context cited by Crystal this would be completely inappropriate. If the accentuation rules were revised so as to produce the result above, they would make a very large number of incorrect predictions.

If we were to demonstrate that the first syllable of *Japanese* in the utterance transcribed by Crystal was less prominent than the third, he would respond that it was "exceptional from the point of view of tonicity" (p. 219). It is not clear whether Crystal intends this as an empirical claim or not. It is one thing to represent a type of utterance as exceptional, it is quite another to show that it is used or perceived in an exceptional way. Crystal makes no attempt to argue, for instance, that native speakers recognize this sort of phrase as unusual. There is no doubt that utterances occur whose accentuation is exceptional. A not uncommon British pronunciation of *rather* is a fairly clear example. But this is, as Lewis (1970:35) says, "an extravagant piece of tonetic slang as far from the mainstream of the language as the position of the adverb in *fan-flipping-tastic*." Native speakers recognize that this way of saying *rather* is exceptional. There is no evidence to suggest that they have similar intuitions about rise-plus-fall utterances of *Japanese*.

If we suppose that *is a Japanese* constitutes a single phrase associated with the tonal morpheme (3)(1)31 (the Long Jump), the normal operation of our accentuation rules will give the correct results. Since *Japanese* is the first focussed word, the first strong syllable, *Jap*, will be made [+accent]. Since *Japanese* is also the last focussed word, its underlyingly [+accent] syllable, *nese*, will be made [+nucleus]. The accent 1-tone is attached to *Jap*, the nucleus 3-tone to *nese*. An optional assimilation rule may associate the nucleus tone with the prenuclear syllable as well.

9. is a 'Japa'nese | |
 3 1 3 1

Crystal's response to the existence of utterances which can be interpreted as either A or B is to propose a third category, C, which will include all the ambiguous examples. This sort of thing can go on indefinitely. Items which are ambiguous between A and C could be classified as instances of D, and so on. The fall-plus-rise is a classic *tertium quid*, which surely cries out for the application of Ockham's razor. If we followed Crystal's example, utterances of /red/ which could be interpreted as either *red* or *read* would lead us to add yet another word to the English lexicon-- *redd*, perhaps.

2. Reasons for binuclear tunes

I will now look at three reasons which have been or might be advanced to justify the introduction of binuclear tunes, and show that none of them are particularly convincing or cogent. Section 2.1 is concerned mainly with Crystal's claim that there is no phrase boundary between some nuclear syllables. I consider in section 2.2 Halliday's claim that there cannot be what he calls a pretonic before the second nuclear syllable of a binuclear phrase. In 2.3 I take up the argument that a binuclear analysis captures a linguistically significant generalization about syntactic structure and phonological phrase boundaries.

2.1. No boundary

I will begin by considering Crystal's reasons for calling these entities single tone units with two nuclear syllables. He promises (1969a:218) to "review the evidence in favour of taking a sequence of kinetic elements as a formal and functional unit", but he does not do so. What he does is give the reader "four phonetic and distributional characteristics" of compound tone units (p. 218):

- i) "The kinetic tones must display an 'endocentric' relationship";
- ii) "There must be no evidence of a tone unit boundary between the tones";
- iii) "One element of the compound tone must be more prominent than the other The phonetically dominant element is usually the first";
- iv) "The second [kinetic element] is the major functional element".

These are evidently criteria to be used for determining whether or not two successive nuclear tones (and nuclear syllables) are included in a single tone unit. Apart from the second, they appear to be arbitrary. I will comment briefly on the others before taking up the second in more detail.

The first criterion is not independent of the others, but directly consequent on the second, as we shall see below. The third criterion does not look particularly useful. In one sense of *prominent* all nuclear syllables are equally prominent. Clearly Crystal is not using the word in this sense. If we understand it in a more phonetic sense, it is not clear whether any two nuclear syllables are **equally** prominent, as they would have to be to be excluded by this criterion. The appearance of the word "usually" means that we cannot exclude sequences in which the second nuclear syllable is more prominent than the first. As for the fourth criterion, it is difficult to see how it could be applied, without criteria for deciding which of two successive nuclear tones is "the major functional element"; Crystal does not provide such criteria.

The second criterion inheres in the concept of the tone unit. If a tone unit boundary coincides with the end of every tone unit, it is not possible for something with such a boundary inside it to be a tone unit, compound or otherwise. To apply this criterion Crystal requires criteria for identifying a tone unit boundary. This is indicated, he says, by two phonetic factors: (a) a pitch change, and (b) "junctural features", especially "a very slight pause" (p. 206). He calls these "phonetic factors" in one paragraph and "phonological criteria" in the next. I will call them phonetic criteria. I will now show that both of them fail.

2.1.1. Pitch change criterion The first phonetic criterion is (pp. 205-6)

a perceivable pitch-change, either stepping up or stepping down, depending on the direction of nuclear tone movement-- if falling, then step up; if rising, then step down; if level, either, depending on its relative height. This is due to the fact that the onset of each tone unit in a speaker's utterance is at more or less the same pitch-level.

If this pitch change is rightly attributed to the onset, then it marks neither the end of one tone unit nor the beginning of the next. Crystal defines the onset

syllable as the first stressed syllable of a tone unit, adding that it is usually pitch-prominent (p. 226). Sometimes it is the first syllable of the tone unit; sometimes it is not. The pitch change associated with it would indicate a tone unit boundary only when the onset syllable was initial in the tone unit. If the second of two successive tone units has a prehead consisting of one or more stressed syllables, then the boundary precedes the prehead, and is not marked by the step up or down at the onset syllable.

A tone unit boundary need not, therefore, be marked by a pitch change. Crystal recognizes this, but in a way which confuses the question rather than clarifying it (pp. 206-7):

If a tone unit with a rising tone and a tail of more than one syllable precedes a tone unit with a relatively long prehead, then the series of low unstressed syllables could theoretically produce a sequence which would be phonetically impossible to split. . . . In such a circumstance there would be no alternative but to have recourse to grammatical or semantic criteria to place the boundary.

First of all, it is impossible to make sense of this without assuming that either "rising" or "low" is a mistake: the syllables forming the tail of a rising tone will not be low. I will assume that "rising" should be replaced by "falling". Second, it is not clear why the prehead of the second tone unit needs to be "relatively long". Crystal does not explain how the length of the prehead is relevant. Perhaps this qualification is introduced in order to reduce the number of utterances which are acknowledged to be problematic. This brings us to the third point. Crystal states that utterances of this type are exceptions which he has to "think up" (p. 206), but it is apparently normal to have sequences of low syllables between the nuclear syllables of fall-plus-rise compound tone units, and these are by no means unusual, accounting for 7.7% of the tone units in his corpus. Since Crystal allows the possibility of placing a boundary on grammatical or semantic criteria, why did he not do so in those utterances?

Even when the second of two successive tone units does not have a prehead, i.e. when the onset syllable is the first syllable, the tone unit boundary need not be marked by a change in pitch. This is because the onset syllable can be low. If it is, and if the final syllable of the preceding tone unit is also low-- as it will be if it is part of the tail of a falling nuclear tone-- then the step up in

pitch supposed to be diagnostic of a tone unit boundary may be absent. Crystal recognizes that low onsets "may cause ambiguity", saying that in such cases "one must have recourse to the second phonetic criterion" (p. 206). In my discussion below of the second criterion I will show that Crystal does not use it in the way he recommends.

Fall-plus-rise utterances are represented as single tone units because there is no phonetic evidence of a tone unit boundary between the nuclear tones. For instance, there is no step up in pitch between the nuclear syllables. This we may grant. But we have seen that the pitch criterion is inadequate: there are tone unit boundaries which are not marked by a step up or down in pitch. Before we turn to the second criterion, we should note that the first one fails in the opposite way as well. There are steps up and down which do not mark TUBs. For example, the rise of a fall-rise nuclear tone commonly-- if not always-- occurs on the final syllable of the tone unit. O'Connor and Arnold (1973:13-16) make it clear that in Southern British English this rise is manifested as a step up in pitch when the final syllable is weak. Crystal is silent on this-- with good reason. He claims that a step up in pitch after a falling nuclear syllable marks the beginning of another tone unit, i.e. that there is a boundary before the higher syllable. O'Connor and Arnold claim-- and Crystal does not explicitly contradict this-- that a step up in pitch after a falling nuclear syllable sometimes changes the nucleus type from a fall to a fall-rise, i.e. that there is a boundary after the higher syllable. It is not possible to decide on phonetic grounds whether the syllable at mid pitch belongs in the same tone unit as the preceding nuclear syllable or in the next one.

2.1.2. Pause criterion The second phonetic criterion of a tone unit boundary is a pause of some sort. It is clear from Crystal's discussion that this criterion is not reliable, either. Like the pitch change, the pause is neither necessary nor sufficient to identify a boundary. On the one hand, Crystal restricts his discussion of "regular definable phonological boundaries" to "not too hurried" speech (p. 205). I take this to mean that TUBs in more hurried speech may not be marked by pauses. On the other hand, there can be pauses which do not mark TUBs, even if the pause is followed by a nuclear syllable before the next pause. This much is clear from Crystal's statement that in his data pauses were

very rare (3 in 500) between the nuclear syllables of compound tone units (p. 206). That is, he analyzed some stretches of utterance containing two nuclear syllables as single (compound) tone units in spite of the presence of phonetic evidence of a tone unit boundary between those syllables. If he had taken the pause as criterial in the absence of a step up or down in pitch between two nuclear syllables, as he himself recommends on the same page, these rare compound tone units containing pauses would have been transcribed as two simple tone units. I also find it difficult to believe that such pauses are in general as rare as Crystal makes out. We do not have to go beyond the records of the Survey of English Usage to find stretches of speech in which they are more frequent. The following excerpt contains two fall-plus-rise tone units, and both have pauses between the nuclear syllables.

I've only | rɛəd . θri: or fɔ: | | at the | mɒmənt | | I | dɒn't θɪŋk ðere's .
 anything . pɑ:rtɪkjʊləli ʌt'stændɪŋ əbaʊt - ʌŋɡəs wɪlsən | | at | ɔ:l | |

(S.3.5a.11)

(In this transcription a full stop indicates a brief pause and a hyphen a "unit pause".)

2.1.3. Phonetic criteria We have seen that Crystal's reason for having compound tone units is not a very good one. Its weakness demonstrates the difficulty of providing a phonetic characterization of phonological phrase boundaries. It seems worthwhile at this point to consider why Crystal attempts to provide phonetic criteria for identifying TUBs. Without such criteria, phoneticians would have difficulty deciding where to locate the boundary in some utterances, and different transcribers might put it in different places, with the result that phonetically identical utterances could have different phonological representations. Declaring that there is no boundary between certain sequences of falling and rising contours allows one to avoid this problem. But it is only a problem if our goal is correct transcription, i.e. if we wish to go unambiguously from sounds to symbols representing a phonological analysis. Such a goal is desirable if we "want our intonational information to be used in the description of grammatical contrasts", as Crystal does (p. 205). It appears that Crystal accepts the following condition:

A phonological description must include instructions for inferring the proper phonological representation of any speech event, without recourse to information not contained in the physical signal.

This is Halle's (1959:21) formulation. Halle shows that acceptance of this condition results in phonological representations which are more complicated than necessary, and unenlightening. This condition seems to have been imposed in order to make linguistics scientific. It is certainly true that physical scientists employ analytic procedures, but as Halle (1959:23-4) points out,

the theoretical constructs which make up the representations discovered by the different types of analysis are . . . postulated without regard for the procedures whereby they can be discovered in the data. Thus . . . it is inconceivable that chemistry would establish substances that can be identified by visual inspection as a category distinct from substances that require more elaborate techniques for their identification.

Linguists who insist that phonological units must be identifiable from acoustic information alone are setting up theoretical entities on a comparable basis. The condition that phonological units be identifiable from acoustic information alone is associated with linguistic theories which require that the units of each hierarchical level be described only in terms of units of the level immediately below it, the phonetic level being below the phonological, and the phonological below the grammatical. This explains why Crystal wishes to use intonational (i.e. phonological) information in the description of grammatical contrasts, and not vice versa.

The absence of phonetic signs of a tone unit boundary between some nuclear syllables leads Crystal to abandon the mononuclear hypothesis. If we do not accept the above condition on phonological descriptions, however, the fact that there are tone unit boundaries which cannot be identified from characteristics of the physical signal is not a problem. Thus Crystal's principal reason for establishing compound tone units has no force for us. It is interesting that Crystal does not ask whether linguistically untrained native speakers of English hear a boundary in compound tone units. Would they have the same difficulty as phoneticians in locating a TUB in such utterances? Is it because he thinks that it is not his business to account for such intuitions that he does not ask? Apparently not. Immediately following his discussion of tone unit boundary identification he says "the presence of a nucleus is what

accounts for our intuition of 'completeness' at the end of the unit"(p. 207). Elsewhere, moreover, he says explicitly that his criterion for establishing prosodic features depends on the intuitions of "a linguistically untrained group of native English speakers" (p. 127). And yet the establishment of compound tone units seems to be a consequence of certain perceived problems of transcribing linguists-- linguistically *trained* native speakers.

2.2. No pretonic

Halliday (1967) gives a different reason for recognizing what he calls double tonic tone groups. He identifies two types: 13 and 53. The first of these, in which tone 1(falling) is followed by tone 3 (low rising), is the equivalent of Crystal's fall-plus-rise. The other is the same as a rise-fall-plus-rise. Since the argument and refutation apply equally well to both, we can restrict our attention to the first type. According to Halliday (1967:13n4), "such tone groups . . . are regarded as single tone groups with double tonic, rather than as sequences of two tone groups" because "it is not possible for a pretonic to tone 3 to occur here following tone 1". This calls for some explication. Halliday's pretonic begins at what he calls the first **salient** (=strong) syllable of the tone group (Crystal's onset syllable) and extends up to the beginning of the **tonic**, which begins with the **tonic** (=nuclear) syllable. The tonic is obligatory, the pretonic optional. There are different kinds of pretonic. For example, a tone 3 tonic may have a mid pretonic, or a low pretonic (or no pretonic).

Presumably Halliday means that when a low rise follows a fall in the same clause, a pretonic never intervenes. Since the syllables following the tonic syllable in a tone 1 (falling) tonic are always low, Halliday is able to claim that all the low-pitched syllables between the tone 1 tonic syllable and the tone 3 (low rising) tonic syllable belong to the first tonic. This claim is actually an empty one, because one might also claim that some of these low-pitched syllables belong to a low pretonic to the tone 3 tonic, and Halliday has no way of showing this counterclaim to be incorrect. This is because the pretonic is defined with reference to the tone group boundary. Since the pretonic is said to begin at the first salient syllable of a tone group, identifying a pretonic depends on identifying a rhythmically strong syllable which precedes a tonic syllable and follows a tone group boundary. With such a characterization of the pretonic, it

makes no sense to use the absence of a pretonic as evidence of the fact that there is no tone group boundary preceding it.

Within the framework outlined in the present work, we can put Halliday's claim in a form which approaches testability: there are no accented syllables between the nuclear syllables of compound tone units. This becomes testable if we focus on utterances which might be analyzed as fall-plus-rise tone units. These would have no pauses and no high syllables between the two nuclear syllables.

Like Halliday, I distinguish between rhythmically strong and weak syllables. His salient syllables correspond for the most part to my strong syllables. Unlike Halliday, I make a distinction between accented and unaccented strong syllables. In my terms the first syllable of a pretonic is non-nuclear accented: [-nucleus, +accent, +strength]; the first syllable of a tonic is nuclear: [+nucleus, +accent, +strength]. To test the above hypothesis we require a fall-plus-rise utterance with at least two strong syllables between the nuclear syllables. If one of these strong syllables is consistently judged to be more prominent than the other(s), then according to our theory it must be accented. This result would falsify our version of Halliday's hypothesis, and, if we assume that there are no postnuclear accented syllables, falsify the claim that the utterance was a single phonological phrase.

In my experience people find it easiest to judge the relative prominence of non-nuclear syllables if they are in the same word. We can meet this requirement by putting in the crucial part of our test utterance a word which normally has two strong syllables, like *Japanese*. We might use the following sentence:

10. I told you the Japanese would reject it.

Recall that in a word like *Japanese* a nuclear accent must go on the last syllable (leaving aside metalinguistic utterances). A non-nuclear accent could go on the first syllable. Suppose we present a fall-plus-rise utterance of the sentence above to native speakers of English, restricting our attention to those who take *ject* to be nuclear, i.e. more prominent than any syllable of *Japanese*. If some of these listeners say that the first syllable of *Japanese* is more prominent than the third, we can take this as a falsification of Halliday's claim about the forbidden pretonic. We can also take it to be a reason for considering this sort of utterance

to consist of two phonological phrases.

2.3. Phrasing and syntax

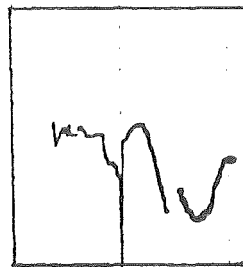
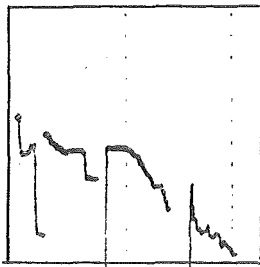
Finally, I would like to consider a kind of argument for binuclear tunes and phrases which does not appear to be offered by any of the linguists whose proposals I have discussed. Indeed, I have not found any explicit statements of it, but it is implicit, I think, in a lot of work on English intonation. The argument is that recognizing binuclear phrases enables us to make certain generalizations about the relationship between intonation, or phrasing, and syntax. In particular, one can say that there is never a phrase boundary immediately before non-initial vocatives. And the generalization is broader than that, because it holds for items other than names. Bing (1980:21) calls items which behave like this Class 0 Expressions. Some of these look like adverbs (*fortunately*), some have the structure of prepositional phrases (*of course*), and some have the structure of sentences (*I'm afraid, you know, I gather*), but most of them appear to function in the manner of what Quirk and Greenbaum (1973:242) call "disjuncts"; other linguists use the less specific term "sentence adverbs". I shall use vocatives for this discussion.

Consider, for example, the sentence *What are you doing, Tony?* If we consider the pitch contour of the vocative, there are two common alternatives. Both syllables may be low, as in 11 below, or the pitch may rise in the final syllable, as in 12.

260 Hz

130 Hz

65 Hz



11. What are you dōing Tony | |

12. What are you dōing Tóny | |

In a binuclear theory, 11 contains one nuclear syllable, and 12 contains two, but both contain only one phonological phrase. Thus one can say that final vocatives are never preceded by a phonological phrase boundary. Obviously this is an attractive generalization. Its truth appears to be assumed by Miller

and Tench (1982:82), who contrast Hausa, where "Vocatives consistently have separate intonation units", with "English, where final vocatives are included in the foregoing unit". In order to make this generalization, they must accept the analysis of the fall-plus-rise as a single phonological phrase. It might therefore be argued that this analysis is supported by its usefulness in capturing this significant generalization about syntax and intonation in English.

This sort of argument is not very persuasive. Even if we accept that there is a generalization to be made here, it is far from obvious that this is the right one. It is certainly not obvious to Bing (1980), who is similarly interested in predicting intonational phrasing from syntactic structure. Bing's generalization about the phrasing of final vocatives is exactly the reverse of the one above. She argues that final vocatives are never "extensions of previous contours" (p. 21): "Vocatives, and all Class O expressions, are obligatorily separated from the rest of the sentence by phrase boundaries" (p. 25). I will take up Bing's claims at greater length in the following section. I introduce them here simply to show that there are other ways of describing the phrasing of final vocatives. In the final section of this chapter I will report an experiment which shows that neither of these generalizations is correct.

Further, as Miller and Tench make clear, their generalization does not cover initial vocatives in English, since they clearly do have "separate intonation units". There are going to be exceptions whether we have binuclear phrases or not. In a system like ours without binuclear phrases, final vocatives are exceptional because they may be included in the same phrase as the preceding words, but initial vocatives are not. In a binuclear theory, all final vocatives are consistently phrased, at the expense of making the phrasing of initial vocatives look exceptional. In short, the behaviour of vocatives and other Class O expressions does not clearly support either hypothesis.

3. Fall-plus-rise as two morphemes

In the concluding sections of this chapter, we turn from bad reasons for abandoning the mononuclear tune hypothesis to good reasons for maintaining it. Since the case for binuclear tunes rests heavily on the fall-plus-rise, I will restrict my attention to this alleged tonal morpheme. I have claimed above

that the typical fall-plus-rise utterance should be represented as two phrases, each carrying a mononuclear tonal morpheme. The time has come to support this claim by showing that such a representation, which is of course consistent with both mononuclear hypotheses, makes correct predictions where alternative representations do not. In section 3.1 we see that the claim that such utterances contain two nucleus tones makes correct predictions about the phonetics of their pitch contours. In 3.2 I show that the claim that such utterances contain two phonological phrases makes correct predictions about the way native speakers perceive them.

3.1. Second nucleus tone

We have been assuming that in order to preserve the mononuclear hypothesis, we must claim that the High Dive, or fall-plus-rise, utterances contain two phonological phrases and two nuclear syllables. There is, however, another possibility. Perhaps such utterances contain only one phonological phrase and one nuclear syllable. This appears to be the view held by Firbas (1980), and Crystal's argument-- discussed above-- against identifying the fall-plus-rise and the fall-rise suggests that others have also taken this view. On the other side of the Atlantic, Bing (1980) agrees that final vocatives and sentence adverbs do not contain nuclear syllables, or-- to report her view more accurately-- that there is no prominence tone attached to such expressions. In fact this solution is incorrect. I have not yet demonstrated this, and neither have any of the linguists who claim that these utterances contain two nuclear syllables or tones. In this section I will repair this omission by showing why the rising tune attached to final vocatives must contain what I call a nucleus tone.

Before proceeding with this demonstration, I should point out that Bing's argument against a nucleus tone is not motivated by a desire to preserve the mononuclear hypothesis. She claims that so-called fall-plus-rise utterances consist of two phonological phrases. Each of these is associated with a contour (tonal morpheme), but only the first contour contains a prominence tone (nucleus tone). (Since Bing's tones are dynamic, her prominence tones are not exactly like my nucleus tones, but it does no harm to equate them in this discussion.) Thus Bing hypothesizes that at least one tune lacks a nucleus tone,

and that such a tune is carried by vocatives and other Class O expressions, as she calls them, in sentence-medial and sentence-final position. In a sense, Bing belongs in the next chapter, where I take up the question of anuclear tunes, but it is also appropriate to introduce her hypothesis here, especially since the data to be presented in the last part of this chapter demonstrate the inadequacy of the boundary aspect of her hypothesis. At the moment, though, we are concerned with her claim that there is only one nucleus tone in the tonal morpheme or morphemes associated with an utterance like *Take the money, Charlie*.

Consider the following sentences, both of which may be spoken with the intonation we are investigating.

13. Good morning, Mitsou. 14. Good morning, Maxine.

In both names the final syllable is strong, but that of *Mitsou* is underlyingly [-accent], while that of *Maxine* is [+accent]. If *Mitsou* contains a nuclear syllable it will normally be the first one. In *Maxine*, the second will be nuclear.

Good 'morning | | 'Mit'sou | | Good 'morning | | 'Max'ine | |

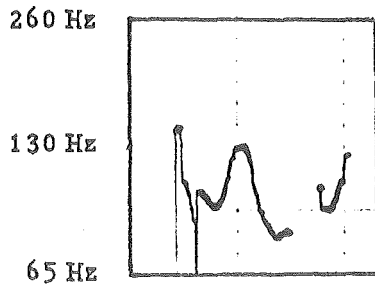
The hypothesis that the tune or subtone in question does not contain a nucleus tone predicts that the location of the nuclear syllable will not affect the distribution of the constituent tones of the tonal morpheme. My hypothesis, on the other hand, predicts that this difference will affect the pitch contour in a particular way. This prediction is simple and precise. In both utterances there is an end tone, say a 3-tone, attached to the final syllable. In *Maxine*, there is also a 1-tone attached to the final syllable; in *Mitsou*, there is not.

Good 'morning | | 'Mit'sou | | Good 'morning | | 'Max'ine | |

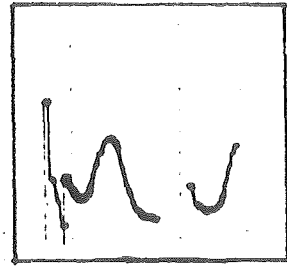
1 3
 | |
 1 3

1 1 3
 | |
 1 3

Thus we predict that in *Maxine*, but not *Mitsou*, the beginning of the second syllable must be low. The beginning of the second vowel of *Mitsou* may be low, because rises in pitch require time, but since it is not phonologically low, it need not be so phonetically. The following traces suggest that this prediction is correct. They falsify the hypothesis that there is no nucleus tone in the tune attached to these words.



13. Good morning, Mitsou.



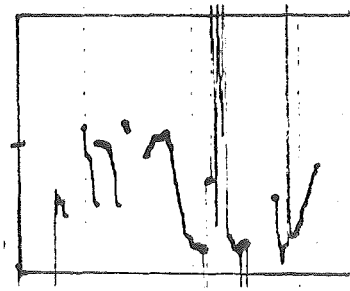
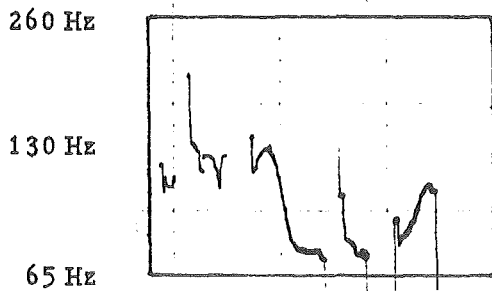
14. Good morning, Maxine.

Lest someone wish to attribute this difference to the different underlying specifications of the final vowels of these names, or some other factor, I will provide another pair of examples, in which each utterance contains the same words.

15. I couldn't see | | Michael's friend | |

16. I couldn't see | | Michael's friend | |

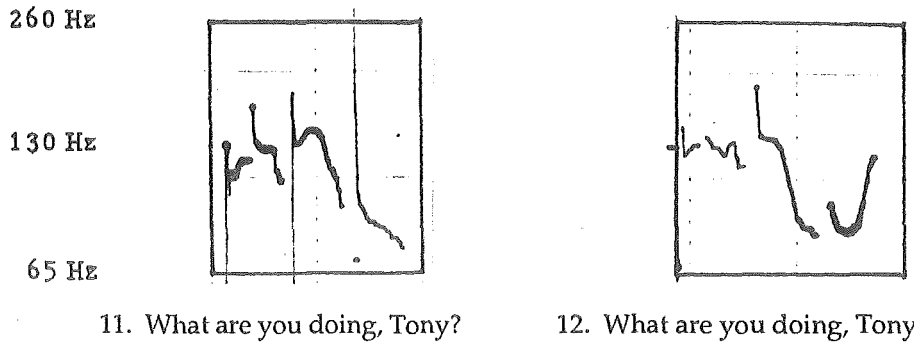
In 15 the first syllable of *Michael* is nuclear, and *friend* is not; in 16 *friend* is nuclear. *Friend* is strong in both versions. My hypothesis predicts that the pitch at the beginning of *friend* will not-- or more precisely, need not-- be low in 15, and that it will be low in 16. Again, this prediction looks to be correct.



15. I couldn't see | | Michael's friend | | 16. I couldn't see | | Michael's friend | |

3.2. Boundary between the nucleus tones

We can now turn to the demonstration that a mononuclear theory is not only simpler than one which abandons the mononuclear hypothesis; it can also account for some facts about the fall-plus-rise which more complicated theories cannot. To put this another way, it makes correct predictions where other theories make incorrect ones about the way native speakers perceive two ways of saying *What are you doing, Tony?*



There are four ways of dealing with this pair of utterances. The first and the last are consistent with the mononuclear hypothesis. If I understand Firbas (1980) correctly, he would analyze both 11 and 12 as one phrase, but would recognize only *do* as nuclear in each one. The second alternative is Bing's. For Bing both 11 and 12 contain two phonological phrases and intonation contours. *Do* has her A Contour in both. In 12 *Tony* has her Class O Contour; in 11 it has the Class 0 Contour without the optional final rise. Crystal would analyze both as single phonological phrases, both with a falling nucleus on *do*, and 12 with a rising nucleus on *To*. Finally, my position is that 11 has one and 12 two phonological phrases, each with a nucleus tone. The four possibilities are summarized below.

- | | |
|---|---|
| 11a. What are you doing Tony | 12a. What are you doing Tony <i>Firbas</i> |
| 11b. What are you doing Tony | 12b. What are you doing Tony <i>Bing</i> |
| 11c. What are you doing Tony | 12c. What are you doing Tony <i>Crystal</i> |
| 11d. What are you doing Tony | 12d. What are you doing Tony <i>DH</i> |

My analysis predicts that native speakers will hear a boundary in 12 but not in 11. The other three analyses do not predict such a distinction. Getting access to people's intuitions can be problematical. We cannot very well ask them whether or not they hear a phonological phrase boundary. We might ask them whether or not they hear a pause, but I did not do so, for two reasons. First, I was not in a position to control the length of a pause precisely. Secondly, such an approach is too blunt. It seems that asking people about their reactions is not as effective as getting their responses without them knowing

they have responded. It also saves a lot of *don't knows* obscuring the results. I did something subtler and simpler, based on the frequent correspondence between phrase boundaries in spoken English and punctuation marks in written English.

A group of native speakers of English heard the sentence *What are you doing, Tony?* spoken three times and were asked simply to write it down. Version 11 was presented to half of the subjects, and 12 to the other half. In speaking these, I took care not to pause at all. My representation predicts a positive correlation between the presence of a final rise and the presence of a comma in the subjects' transcriptions. The alternative representations of the two versions of the sentence predict that there will be no such correlation.

In most sentences spoken with the intonation we are investigating, a comma is either forbidden or obligatory according to the standard grammatically-based rules of English punctuation. I chose to use a sentence in which a comma is obligatory. I expected that this would not prevent my subjects omitting it, even though they were all enrolled in first-year university English courses, but the extent to which they omitted it is surprising. The comma appeared in only 34 of 87 transcriptions: less than 40%. The results are tabulated below. Forty-four subjects were presented with the sentence without a final rise; 11 of these (25%) wrote it with a comma after *doing*. Forty-three subjects were presented with the sentence with a final rise; 23 of these (54%) wrote it with a comma.

	No	Comma	Total
No Rise (11)	33	11	44
Rise (12)	<u>20</u>	<u>23</u>	<u>43</u>
Total	53	34	87

The correlation between the presence of a rise and the presence of a comma is very clear. We can be confident that this is not a chance correlation. According to the chi-square test of statistical significance, the probability that the relationship is a chance or random one is less than 0.01. Obviously the presence of a comma is not entirely determined by the intonation. If it were, then the study of language would be boring. It is clear for a start that the

standard rules of English punctuation are a mystery to most of my subjects. The presence of commas in the transcriptions of 11 we can explain by supposing that these people know the punctuation rules, which require commas before vocatives. The absence of commas in half of the transcriptions of 12 is not so easy to explain. The low number of commas may be due to the fact that I took care not to pause at all before the vocative in the test utterances. Perhaps the producers of these transcriptions insist on hearing silence before writing a punctuation mark. The fact that our mononuclear theory does not by itself account for the presence of commas in some transcriptions of 11, or the absence of commas in some transcriptions of 12 need not concern us, since I know of no other theory which can explain these data. The point is that our theory predicts that the final rise will affect the transcriptions in the way that it does; the competing theories do not.

I began this chapter by outlining some proposals for binuclear tunes which would require us to give up the hypothesis that each tonal morpheme contains one and only one nucleus tone. Then I reviewed some of the arguments in favour of complicating the theory of English intonation in this way, and argued that they were not compelling. Having demonstrated that the complications are unnecessary, I showed, finally, that such theories are also inadequate compared to a simple mononuclear theory when it comes to accounting for certain phonetic and psychological aspects of the so-called fall-plus-rise contour. Far from forcing us to abandon the mononuclear tune hypothesis, the fall-plus-rise provides evidence of its superiority.

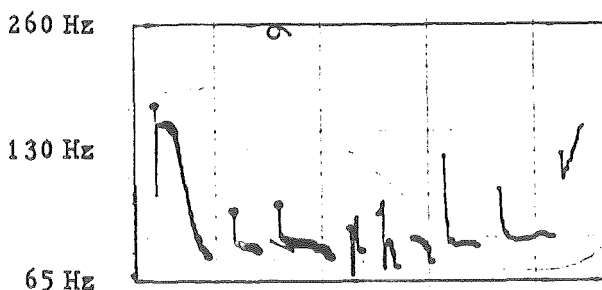
Chapter 8

Anuclear Tunes

We have repelled assaults on the mononuclear tune hypothesis from one direction in the shape of tunes which appear to contain more than one nucleus tone. This chapter is concerned with an assault from the opposite front, in the shape of anuclear tunes-- tonal morphemes which appear to have no nucleus tone. This assault will be more successful than those dealt with above; the mononuclear tune hypothesis will not survive it intact. A loss on this front, however, is not as damaging as a loss would have been on the binuclear front. At the end of the chapter I will briefly discuss three factors which mitigate the seriousness of this loss.

The tonal morpheme we deal with in this chapter is the one which Liberman and Sag (1974) label the Contradiction Contour. This is not the only anuclear tune which has been proposed. For example, Bing (1979) claims that her Class 0 Contour lacks what I would call a nucleus tone. I presented some evidence against this claim at the end of the preceding chapter, and I will not repeat the argument here.

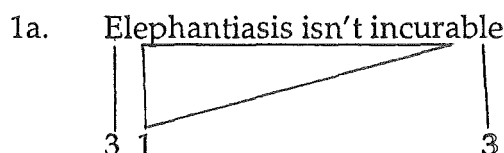
Liberman and Sag (1974:420) introduce the tune in question with the sentence *Elephantiasis isn't incurable*, spoken in the following way.



1. Elephantiasis isn't incurable

According to Liberman and Sag, the pitch contour of such phrases is high in the first syllable and in the last, and relatively low in between. They propose that this contour is a "holistic unit". Insofar as this equates to a claim that it is a single morpheme, I would agree. They also assume that, in my terms, it contains no nucleus tone, though they offer very little evidence to support this claim. I will argue that this, too, is correct. I will disagree, however, with most of the other claims they make about this tune.

In the framework of this study, Liberman and Sag's phonological analysis of this tonal morpheme can be fairly represented as 313. All three tones are obligatory. The ITA rules will attach the end 3-tone to the last syllable of a phrase, and the WF rules will attach the initial 3-tone to the first syllable, and the 1-tone to all the intervening syllables. This gives us the following representation for 1, which correctly predicts the pitch contour shown above.



(Actually, it is possible that the final tone is higher than the first-- a 4-tone rather than a 3-tone-- but I will not be discussing this detail of the representation here.) Liberman and Sag make-- explicitly and otherwise-- three claims about the first 3-tone which I will take issue with. First, they claim that it is "uncorrelated with word stress". They also assume that it is obligatory and that it is the first tone of the morpheme. I will argue, contrariwise, (i) that it is an accent tone, (ii) that it is optional, and (iii) that it is not morpheme-initial. In short, I will propose that this tonal morpheme be represented as [33]13. The accent tone within square brackets is attached to an accented syllable, but unlike ordinary accent tones, need not appear whenever the associated phrase contains such a syllable. I will also take issue with Liberman and Sag's (1974:421) claim that "this contour is appropriate ... just when the speaker is using the utterance which bears it to contradict".

The following discussion will not, however, be devoted entirely to Liberman and Sag's proposals about the contradiction contour. This chapter is not organized strictly according to the hypotheses outlined in the preceding paragraph, because I wish to consider some other analyses of this tune. In

particular, I will examine some of the alternative representations embodied in transcriptions by British writers on intonation, and show why they do not work. I will begin with the hypothesis that the first 3-tone is a nucleus tone. In section 2, I point out the weaknesses of the hypothesis that the first 3-tone is followed by an accent 1-tone. The hypothesis that the tune contains no accent tone is attacked in section 3, and the claim that it contains a nucleus tone in section 4. I take issue in section 5 with the hypothesis that this tonal morpheme contains no optional tones. In the sixth and final section I argue that the term "contradiction contour" is inappropriate, because the same tune occurs with utterances which have a quite different discourse function.

1. Status of initial tone

Let us start by considering transcriptions in which the high first syllable is represented as linked to a nucleus tone. Crystal (1969a:273) offers the following as a transcription of an utterance "signalling opposition", that is, an utterance used to contradict.

2. |Yès||it|ís||

This representation appears to be the equivalent of a Drop followed by a High Bounce, which takes the following form in our system.

2a. 'Yes || it 'is ||
 \wedge \wedge
 3 1 1 45

This would be an appropriate way of saying the two sentences *Yes. It is?*, which sounds quite different from an argumentative *Yes it is!* A somewhat better solution would be to analyze the tune of *it is* as a High Level: (1)(3)4. This would at least distinguish the utterance signalling opposition from *Yes. It is?*

3a. 'Yes || it 'is ||
 \wedge \wedge
 3 1 1 4

This is roughly equivalent to the following in Crystal's system:

3. |Yès||it|ís||

It is not impossible that this is what Crystal intended, and that the rising tone mark above *is* in his transcription is a misprint. Another possibility is that the second tone-unit was intended to have a low onset, indicated by a small

majuscule L before *is*, as follows.

4. |Yès||it|_Lís||

In our terms this means that the second phrase carries the Take Off tune (or the Low Bounce). We shall see below that other linguists use the equivalent of this tonal morpheme to represent the contradiction contour.

- 4a. 'Yes||it'ís||
 $\begin{array}{cc} \wedge & \wedge \\ \underline{3} & 1 & 1 & \underline{1} & 3 \end{array}$

Both of these alternatives are better than the printed version, but all three are unsatisfactory. There is no warrant for the tone-unit boundary after *yes*. The grammatical structure of this example makes a boundary here plausible, because it precedes a complete clause. It is much less plausible in other utterances signalling opposition, such as *You don't know!*, which Crystal himself transcribes elsewhere (1969:220) as a single tone unit with a falling nuclear tone on the first syllable and a rising one on the last.

5. |Yòu don't knòw|| (so why are you saying you do!)

The same transcription is used in the files of the Survey of English Usage for some of the contradictory utterances of this type in that corpus. In a search of the files I found seven utterances which are used to signal opposition, as Crystal puts it, and appear to carry the contradiction contour. Three of these are transcribed like "You don't know!" above, with a fall plus a rise. (I will deal with the others in due course.)

6. |thìs isn't ríght|| (S.3.3.41)
 7. |nò it dóesn't|| (S.2.10.13)
 8. |nò it's nót|| (S.1.11b.68)

(It is likely that the SEU contains many more samples of the contradiction contour, but finding them is a problem. I was able to find these only because there is a file of utterances containing *no* or *not*.)

The claim implicit in all of the above transcriptions is that the first 3-tone of the contradiction contour is a nucleus tone. This claim is incorrect. Example 1 above makes it clear that this tone is attached to syllables which are not nuclear. In that utterance the first 3-tone is attached to the first syllable of *elephantiasis*, which is [-accent] in the underlying form of the word. Recall that a syllable which is underlyingly [-accent] will become [+nucleus] only if it is

focussed and the [+accent] syllable of the word is not. There is no reason to suppose such a narrow focus for the example under discussion. Since such an utterance is not limited to situations in which the first part of *elephantiasis* is appropriately focussed, we can assume that the first syllable is not nuclear. Since the 3-tone is attached only to this syllable, it cannot be a nucleus tone.

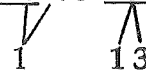
Lieberman and Sag (1974:421) also make this point, but they are not clear about its significance, claiming that the initial fall in pitch is "uncorrelated with word stress". What we have demonstrated is that the first 3-tone of this tonal morpheme is not a nucleus tone. We have not demonstrated that this tone is not an accent tone. While the first syllable of *elephantiasis* is not underlyingly [+accent], it is [+strength] and would normally become [+accent] when it is the first such syllable in a phrase. Neither have we demonstrated that the contradiction contour contains no nucleus tone. It does not follow from the fact that the first tone is not a nucleus tone that none of the others is.

2. Status of 1-tone


Having eliminated the hypothesis that the first 3-tone is a nucleus tone, let us now consider the hypothesis that the first two tones of the contradiction contour are 3 (unmarked) and 1 (an accent tone).

O'Connor and Arnold (1973:58) say that their Take-Off tone group is very common in resentful contradictions. These obviously have the same function as Crystal's utterances signalling opposition, and Lieberman and Sag's contradicting utterances, but they are phonologically not quite the same, having no 3-tone at the beginning.

9. 'No it's 'not.

'No it's 'not


10. You 'can't do 'that.

You 'can't do 'that.


Sometimes, they say, the contradiction sounds deprecatory.

11. I 'don't 'know. | | 'You could have 'done it 'just as 'well. | |

Halliday (1967:25) also notes the use of this intonation, which he describes as a rising tonic (neutral version of Tone 2) with a low pretonic (i.e. head), in

challenging, aggressive, defensive, or indignant contradictions.

12. || --2 I don't I know ||

Halliday's transcription indicates that he considers the first syllable to be stressed rather than the second, which is stressed in O'Connor and Arnold's version.

The high-start contradiction contour described by Liberman and Sag is treated by O'Connor and Arnold (1973:158) as a variant, an emphatic variant, of the Take-Off: "All the relevant drills given above with the tune (LOW PREHEAD +) LOW HEAD + LOW RISE (+ TAIL) can be said with emphasis if the high prehead is used instead of the low prehead". The first syllable of the head, remember, is the first strong syllable in the phrase. In our terms, then, O'Connor and Arnold propose that the contradiction contour has this form:

(3)(1)13.

13. 'You 'can't do 'that. (p. 36)

You 'can't do 'that
 3 1 1 3

There are several problems with this hypothesis. The most obvious is that many phrases bearing the contradiction contour begin with a strong syllable. This is true of all the examples from Crystal and the Survey of English Usage, and in fact the vast majority of such utterances in English. It is clear, then, that phrases like *Yes I have* have "emphatic variants", just as phrases like *You can't do that* (with the first syllable weak) have them. O'Connor and Arnold can provide an emphatic variant for the latter, but not for the former, because it has no prehead to make high. According to their solution, a phrase can be said with emphasis only if it has a prehead. This notion would be suspect even if we did not possess numerous counterexamples.

It must be said that O'Connor and Arnold are aware of the problem. But the example they use in attempting to justify their solution provides more evidence against it (p. 35).

14. °People 'won't eat 'that.

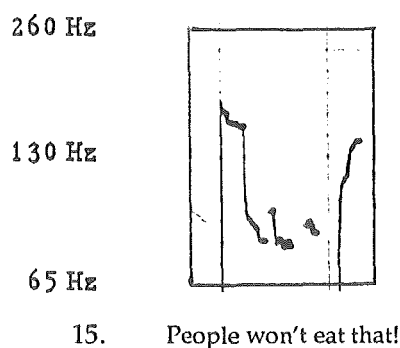
'People 'won't eat 'that
 3 1 1 3

The utterance is clearly a contradicting, or objecting one. They describe it as a Take-Off tone group with a high prehead beginning with a strong syllable.

Strong syllables before the first accented syllable of a phrase are unusual, as O'Connor and Arnold recognize. But the problem here is that their representation makes incorrect predictions, as I shall show directly. Since their solution does not work for their own example, generalizing it to other phrases, such as *Elephantiasis isn't incurable*, is out of the question.

There are two aspects of O'Connor and Arnold's transcription of a contradicting *People won't eat that* which are questionable, to say the least. The first is the relative prominence of the first and third syllables. Since the first syllable is marked as strong but not accented, the transcription embodies a claim that the third syllable, which is accented, is more prominent than the first. This claim seems quite wrong to me, and O'Connor and Arnold offer no evidence to support it. Testing the claim would be relatively easy. Our refutation would be most convincing if we could use a recording by the authors of this utterance, but it is unfortunately not among those in the recording which accompanies their book.

The second questionable claim implicit in their transcription of *People won't eat that* is easier to refute, because the data involved are more phonetic than psychological. According to their representation the second syllable of *people* is phonologically high, i.e. attached to a 3-tone. We cannot doubt that a phonetician could produce such an utterance, but we may doubt that the contradiction contour has this shape. When I produce one, the second syllable is always low, as in the following trace.



The first break in the trace is the [p] at the beginning of the second syllable, which is clearly not high. The pitch is still descending rapidly at the beginning of the vowel, but this pattern is exactly what we should expect if the first

syllable is attached to a 3-tone and the second syllable to a 1-tone.

Perhaps my intonation is odd. This seems unlikely, though, because other linguists' descriptions and transcriptions of the contour also indicate a low second syllable when the first is strong. If they perceived the second syllable to be high, we would not find them indicating a falling nuclear tone on the first syllable. We saw in section 1 that some linguists do this. I conclude that the contradiction contour is not to be represented as a Take Off with a high prehead, that is, as 3113.

3. Location of accent tone

We move on to a second hypothesis which attributes only one nucleus tone to the contradiction contour. In particular, let us consider a representation which differs from O'Connor and Arnold's in that the first of the two 1-tones is not an accent tone: 3113. According to this hypothesis, the tonal morpheme in question contains no accent tone.

This hypothesis makes incorrect predictions about the intonation of phrases like the following, in which the first syllable is clearly weak.

16. Pneumonia isn't incurable!

17. Impossible isn't an adverb!

Given our tone association rules, the representation above predicts that the first syllable will be attached to a 3-tone and the second to a 1-tone. I will assume that *poss* is accented and *ad* nuclear.

Im'possible 'isn't an 'adverb

3 1 1 3

ITA Rules

Im'possible 'isn't an 'adverb

3 1 1 3

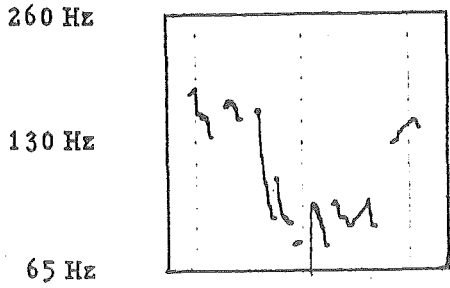
WF Rules

Im'possible 'isn't an 'adverb

3 1 1 3

This is incorrect-- for my dialect, at least. The following record shows that the second syllable is high as well as the first. The third vowel, after the gap

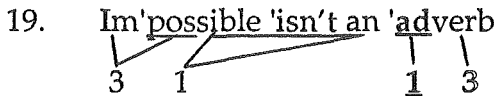
corresponding to the [s], is the first low one.



18. Im'possible 'isn't an 'adverb

Both the first and the second syllables appear to be attached to a 3-tone. Neither is attached to a 1-tone.

What we require, apparently, is something like this:



The 1-tone cannot be attached by an ITA rule, because the leftmost syllable to which it is linked is clearly not accented. Let us suppose instead that the initial 3-tone is an accent tone. This will give us the following derivation.

Im'possible 'isn't an 'adverb

3 1 1 3

ITA rules

Im'possible 'isn't an 'adverb

3 1 1 3

WF Rules

Im'possible 'isn't an 'adverb

3 1 1 3

This correctly predicts that *ible* is low, and that *poss* is high, but it makes no prediction at all about the first syllable. Because we are assuming that tone spreads to the right, not to the left, the WF rules will not attach the first syllable to the accent 3-tone. Rather than tamper with these universal conventions, let us add another 3-tone to the left of the accent 3-tone, thus: 33113. The WF rules will attach this to any syllables before the pre-nuclear accented syllable, and give us a representation which correctly predicts the pitch contour of a

contradicting utterance of *Im'possible 'isn't an 'adverb*.

20. Im'possible 'isn't an 'adverb
 3 3 1 1 3

4. Absence of nucleus tone

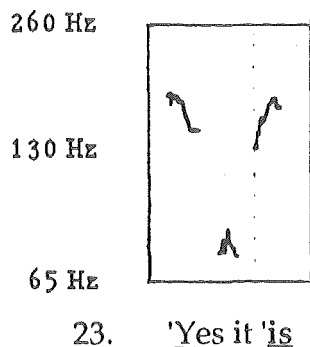
I have not so far questioned the representation of the penultimate tone as a nucleus tone. In this section I will show briefly that neither the penultimate tone nor the final tone is a nucleus tone. We can test the correctness of a representation ending in 13 by examining phrases in which the final syllable is nuclear. Let us consider an utterance of *Yes it is* with a contradiction contour. We can assume that the final syllable is nuclear. If a contradicting *Yes it is* is spoken as a single phrase with the Drop tune, *is* will be nuclear, i.e. attached to both a nucleus 3-tone and a 1-tone.

21. 'Yes it 'is | |
 3 3 1

It is reasonable to assume that changing the tune will not affect the location of the nuclear syllable. If we analyze the contradiction contour as 33113, the same phrase with this tune will be represented as follows.

22. 'Yes it 'is | |
 3 1 1 3

This representation predicts that the final syllable must be high at the end, and low at the beginning. This prediction is incorrect. The following trace shows that the final syllable of such an utterance does not have to be low at the beginning. The last gap in the trace is the voiceless [t] of *it*.



I conclude that the nucleus 1-tone is unnecessary, and omit it: 3313.

I have now demonstrated that the penultimate tone of the contradiction contour is not a nucleus tone. Again, this does not mean that there is no nucleus tone, but we are running out of candidates. Perhaps the final tone is worth consideration. This kind of representation would work for *Yes it is*.

24. 'Yes it 'is | |
 | | |
 3 1 3

Unfortunately it will not work for a great many other contradictions, such as *Elephantiasis isn't incurable*. Here the nuclear syllable cannot be after *cure*, but the pitch does not rise until the final syllable. I can think of no other possibilities for a nucleus tone in the contradiction contour, as long as we assume that all such contours realize a single tonal morpheme. I am reluctant to abandon this assumption, because our brains and people's use of these tunes insist that they are all the same, not different morphemes.

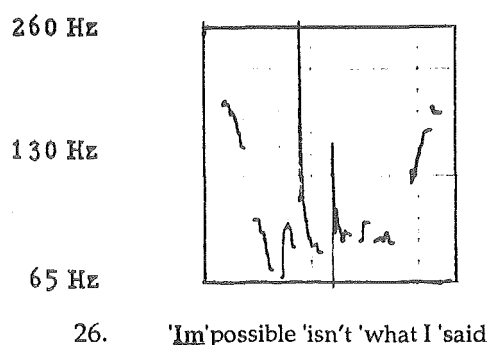
I have now proposed that the tonal morpheme known as the contradiction contour be represented as 3313. Unlike the accent tones of other morphemes, this one is not enclosed in parentheses. I have not put round brackets around it because it is different from an "ordinary" accent tone. The accent tones we have considered up to this point are attached by the ITA rules only to prenuclear accented syllables. If the associated phrase has no accented syllable before the nuclear syllable, such tones remain unattached. The accent tone of the contradiction contour does not remain unattached under these circumstances. In the absence of a prenuclear accented syllable, it is linked to the nuclear accented syllable.

Since it can be difficult to identify the nuclear syllable of contradiction contour phrases, let us take advantage of the fact that the Nuclear Accentuation Rule can accent underlyingly [-accent] syllables, but the Prenuclear Accentuation Rule cannot. Thus the first syllable of *impossible*, for example, may become [+nucleus, +accent], but not [-nucleus, +accent]. So if the first syllable of *impossible* is more prominent than the second, it must be nuclear as well as accented. This will happen when *im*, but not *possible*, is focussed.

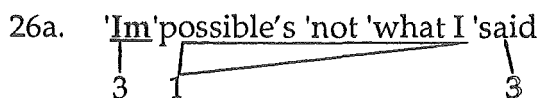
25. 'Im'possible 'isn't 'what I 'said

When this phrase is spoken with the contradiction contour, the first syllable is

high, and the second syllable is low. That is, the second syllable is the first one which is attached to a 1-tone.



This corresponds to the representation our rules produce if the accent 3-tone is attached to the nuclear syllable in the absence of a prenuclear accented syllable.



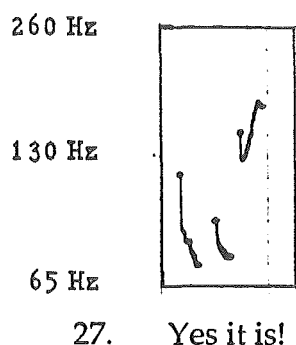
5. Optional tones

By identifying the contradiction contour as holistic, Liberman and Sag appear to claim that all its tones are obligatory. In this section I will take issue with this hypothesis, proposing that the accent 3-tone and the one preceding it are optional. In this way we account for the facts which moved O'Connor and Arnold to treat the contradiction contour as an emphatic variant of the Take-Off tune, which lacks any 3-tone at the beginning.

We may as well begin our discussion by considering the appropriateness of the labels provided by O'Connor and Arnold. Insofar as they suggest that the emphatic or high contradiction contour is less common than the low version, they are not appropriate. The evidence I have seen suggests that the so-called emphatic variant is much more common. Liberman and Sag's failure to mention the low variant points in this direction. More convincing evidence is to be found in the Survey of English Usage. I mentioned earlier that this corpus of spoken English (RP) contains several examples of resentful contradictions. Of the eight I have found, none begins with a low syllable. It is possible, of course, that the speakers represented in the survey are particularly

excitable or forceful, and that the proportion of high variants is abnormal. This is unlikely if there is any truth behind the phlegmatic reputation of speakers of this dialect. We should beware of describing the more common variant as the emphatic one, in the absence of convincing evidence to support this description. I will henceforth refer to the variants as high and low.

O'Connor and Arnold treat the low contradiction contour as an instance of their Take-Off, and the high one as a variant. This cannot be correct, because the Take-Off, which in our framework is (1)(1)13, is clearly different from the low contradiction contour. The penultimate tone of the Take-Off morpheme is a nucleus tone. The penultimate tone of the low contradiction contour, like that of the high variant, is not. We can use the phrase *Yes it is!* again to show this. If the penultimate tone were nuclear, the final syllable would have to start low. It does not, as the following trace shows.



Let us suppose, therefore, that the first two tones of the high contradiction contour are optional, and that the low variant is derived by omitting them. The whole morpheme looks like this: [33]13. I enclose the preaccent tone together with the accent tone, because I am assuming that they cannot be selected independently. Without the optional tones it has this form: 13. The standard tone association rules will link the first tone with the first syllable of the phrase, no matter what its status. The 3-tone will be attached to the final syllable. Since all the preceding syllables are low, we need only a single 1-tone, and there is no need to complicate matters with an accent tone.

I have described the two tones in square brackets as optional. I have used the same word to describe the accent tones in other morphemes, which are enclosed in round brackets. This is something of a misnomer, since the accent

tones in round brackets are obligatory if there is a non-nuclear accented syllable in the phrase. It would be more accurate to say that the round brackets around accent tones prevent the WF rules from attaching them if they are not attached by the ITA rules. The accent tone in the contradiction contour, however, is not obligatory when the phrase contains a prenuclear accented syllable. If it behaved like a standard accent tone, only the high variant could occur if the phrase contained an accented syllable. The fact is that accented syllables do not prevent the use of the low contradiction contour, as its use with 'Yes it 'is, and comparable phrases cited in section 3, makes clear.

Actually, it is not necessary to use two kinds of brackets. The round brackets are certainly useful for teaching and discussion, but they can be omitted, because all accent tones outside of the contradiction contour behave in the same way. Instead of marking such tones, we can simply state that the restriction described above applies to all unmarked accent tones. The significance of the round brackets around preaccent tones is slightly different, because these tones are normally first attached by WF rules. The brackets indicate that these tones are not to be attached by the English-specific well-formedness rule, WF Rule 4. But again, since this restriction seems to hold for all tones which precede the accent tone of a morpheme, the round brackets may be dispensed with. In fact, we can probably dispense with the square brackets in the contradiction contour as well, by assuming that the behaviour of its accent tone (and the one before it) is characteristic of accent tones in tonal morphemes which lack a nucleus tone. Nevertheless, I will continue to represent the contradiction contour as follows, with the first two tones in square brackets: [33]13.

6. Other uses of the contour

Finally, I take up the claim made by Liberman and Sag (1974:421) that "this contour is appropriate . . . just when the speaker is using the utterance which bears it to **contradict**". This is incorrect, because the same tonal morpheme is used with utterances having entirely different functions.

One kind is the *Here you are* which accompanies giving. This can be

heard often in shops at the end of a transaction. Other sentences which can be said with the same tune are *There you are*, *Here you go*, and *There you go*.

The Survey of English Usage contains two examples, spoken by a person giving someone a drink.

28. |hère we áre|| (W.5.2.28)

29. |hère you áre|| (W.5.2.43)

This is transcribed with a fall-plus-rise nucleus, as are most of the Survey instances of contradictions with a high start. O'Connor and Arnold include two utterances of this kind, which they transcribe in the same way.

30. `There you `are. (Your `library obook.) (p. 147)

31. Oh `here you `are. ('Catch hold of `this one.) (p. 233)

A lexically and pragmatically similar kind of utterance which uses the same tune is the *Here we are* that frequently accompanies finding or revealing.

Again the Survey of English Usage contains examples, twice transcribed as above, and once with a single fall-rise nucleus on the first syllable.

O'Connor and Arnold supply examples of yet another type of utterance which can carry this tune. They are used to perform an illocutionary act which to my knowledge has no name. We might call it minimizing. Such utterances can be used in response to apologies, but they are not essentially forgiving.

32. It `doesn't `matter. (p. 236)

33. There's `no real `harm done. (p. 236)

34. `Never `mind. (p. 240)

O'Connor and Arnold offer these as examples of the High Dive, but they are not. Identifying these as instances of the "contradiction contour" explains why they escape a generalization about High Dive phrases which O'Connor and Arnold (1973:85) state as follows:

We [speakers of English] use the High Dive then whenever the first part of a word group contains the most important idea, and the second part an idea of subsidiary importance.

They are apparently referring to the fact that High Dive phrases (granting for the moment that they are single phrases) are very similar in their pragmatics to phrases with the Drop tune and a single nuclear syllable, and that in the

corresponding single-nucleus phrase, it is the first of the two nuclear syllables which is preserved. We saw this in Chapter 7 with the sentence *What are you doing, Tony?* The single-nucleus phrase which corresponds to 35 is 36, rather than 37.

- 35. What are you **doing** Tony
- 36. What are you **doing** Tony
- 37. What are you **doing** Tony

This observation holds for most of O'Connor and Arnold's High Dive phrases. A couple of examples should suffice here:

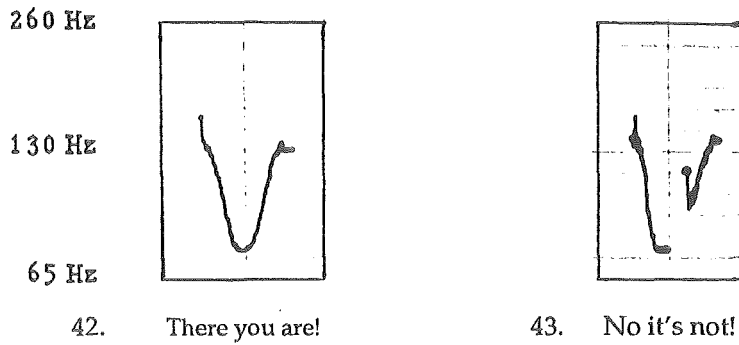
- 38. My `mother ◦came from `there. | |
- 39. He's `mad on ◦apple `pie. | |

The generalization is not valid, however, for minimizing utterances such as 32-34 above. In the corresponding single-nucleus phrase, it is the first nuclear syllable which will be eliminated rather than the second. Thus 34 corresponds not to 40 but to 41.

- 34. `Never `mind
- 40. Never mind
- 41. Never **mind**

Moreover, if the pragmatics are to remain approximately the same, the tonal morpheme cannot be a Drop. This, too, argues against identifying the tune of 34 as the High Dive.

In order to show that the contradiction contour does not occur only with contradicting utterances, I have presented three kinds of utterances which appear to be spoken typically with the same tune. So far I have offered little in the way of support for my claim. It would be tedious to repeat all the phonological arguments given earlier for contradicting utterances. Some of the arguments simply cannot be used for giving and finding utterances like *There you are*. Because these are formulas, we cannot investigate their intonation by replacing them with other sentences having the same effect. What we can do is show that the pitch contour of such an utterance, as recorded by laryngograph, is identical to the contour of a resentful *No it's not!*



Another type of fact which supports our identification of these tunes has already been mentioned. Like contradictions, these utterances frequently occur with the low variant of the contour. The following is a minimizing utterance from O'Connor and Arnold (p. 246):

44. Oh there's 'no real 'harm done.

Once again, we find examples of giving and finding utterances in the Survey of English Usage:

45. here you |áre| | (S.4.6b.8)
 46. |here we áre| | (S.6.2.34)
 47. |_Lhere we áre| | (W.5.2.3)

Another piece of evidence is the inconsistency of this tune's representation in the corpus. As with the contradictions, we find both high and low variants of these utterances transcribed in a variety of ways. This is not because they are spoken with several different tunes, but because neither of them conform to the assumptions behind the transcription system. Since they do not fit into the system, it is not surprising to find different transcribers -- or the same transcriber at different times -- dealing with them in different ways.

In the preceding discussion of the tonal morpheme which Liberman and Sag call the contradiction contour, I have argued that some of the claims they make about it are wrong. After entertaining and dismissing several possible representations, I settled on this form: [33]13. The most obvious oddity is the absence of a nucleus tone. The status of the accent tone is also unusual, since it need not be used, even if the phrase contains an accent tone. And since there is no nucleus tone, it can actually be attached to a syllable which is nuclear, if that syllable happens to be the first accented syllable of the phrase. If my analysis is

correct, we must abandon the mononuclear tune hypothesis. As I suggested at the beginning of this chapter, however, the existence of a morpheme with no nucleus tone is not so disturbing as the existence of a binuclear tune would be. There are three reasons for this.

First of all, the recognition of anuclear tunes, unlike the recognition of binuclear tunes, does not endanger what I have called the mononuclear phrase hypothesis, i.e. the hypothesis that each phonological phrase contains one and only one nuclear syllable. Claiming that a tonal morpheme lacks a nucleus tone does not require us to predict that the phrases to which it is attached do not contain nuclear syllables.

Second, we can still maintain the hypothesis that no tonal morpheme contains more than one nucleus tone. This may be called the Weak Mononuclear Tune Hypothesis, to distinguish it from the other, which we might rename the Strong Mononuclear Tune Hypothesis.

Third, the tune which falsifies the latter is exceptional in other ways. In particular, it contains an accent tone which behaves differently from the others we have encountered. It could be accidental that this tune is exceptional in two different ways, but it is also possible that the two characteristics are related. The Strong Mononuclear Tune Hypothesis would therefore still hold for tunes with ordinary accent tones, those which may not be attached to a nuclear syllable and must occur when a prenuclear accented syllable is present in the associated phonological phrase. We do not yet have to resort to a hypothesis which amounts to a statement that all tunes contain a nucleus tone except those which do not. It is reasonable to expect a tune which is formally exceptional to be functionally exceptional, as well. If we were to find that this anuclear tonal morpheme was used in unusual ways, the discovery would lend support to an analysis which identifies it as formally exceptional. Whether the use of the so-called contradiction contour does set it apart from other English tonal morphemes is a question which calls for further investigation.

References

- Abercrombie, David, D. B. Fry, P. A. D. MacCarthy, N. C. Scott, and J. L. M. Trim (eds). 1964. *In honour of Daniel Jones: papers contributed on the occasion of his eightieth birthday, 12 September 1961*. London: Longman.
- Allen, William Sidney. 1973. *Accent and rhythm*. Cambridge: Cambridge University Press.
- Anderson, Stephen R. 1978. Tone features. In Fromkin 1978: 133-176.
- Anderson, Stephen R. 1985. *Phonology in the twentieth century*. Chicago: Chicago University Press.
- Anshen, Frank. 1978. *Statistics for linguists*. Rowley, Mass.: Newbury House.
- Berry, R., and J. Acheson (eds). 1985. *Regionalism and National Identity: Multi-Disciplinary Essays on Canada, Australia, and New Zealand*. Christchurch, N.Z.: ACSANZ.
- Bing, Janet. 1980. *Aspects of English prosody*. Bloomington: Indiana U. L. C. (1979 Ph.D. thesis, University of Massachusetts.)
- Bloomfield, Leonard. 1933. *Language*. New York: Holt, Rinehart and Winston.
- Bolinger, Dwight L. 1949. Intonation and analysis. *Word* 5: 248-254.
- Bolinger, Dwight L. 1951. Intonation: levels versus configurations. *Word* 7: 199-210.
- Bolinger, Dwight L. 1972a. *Degree Words*. The Hague: Mouton.
- Bolinger, Dwight L. 1972b. Accent is predictable (if you're a mind-reader). *Language* 48: 633-644.
- Bolinger, Dwight L. (ed). 1972c. *Intonation: selected readings*. Harmondsworth: Penguin.
- Bolinger, Dwight L. 1986. *Intonation and its parts: melody in spoken English*. London: Edward Arnold.
- Brazil, David, Malcolm Coulthard, and Catherine Johns. 1980. *Discourse*

- intonation and language teaching*. London: Longman.
- Chomsky, Noam, and Morris Halle. 1968. *The sound pattern of English*. New York: Harper and Row
- Clements, George Nick and Kevin C. Ford. 1979. Kikuyu tone shift and its synchronic consequences. *Linguistic Inquiry* 10: 179-210.
- Clements, George Nick, and Samuel Jay Keyser. 1983. *CV Phonology: A generative theory of the syllable*. Cambridge, Mass: MIT Press.
- Comrie, Bernard. 1981. *The languages of the Soviet Union*. Cambridge: Cambridge University Press.
- Cruttenden, Alan. 1986. *Intonation*. Cambridge: Cambridge University Press.
- Crystal, David. 1969a. *Prosodic systems and intonation in English*. Cambridge: Cambridge University Press.
- Crystal, David. 1969b. A forgotten English tone: an alternative analysis. *Le Maître Phonétique* 132: 34-37.
- Firbas, Jan. 1980. Post-intonation-centre prosodic shade in the modern English clause. In Greenbaum *et al.* 1980.
- Fox, Anthony. 1969. A forgotten English tone. *Le Maître Phonétique* 131: 13-14.
- Fox, Anthony. 1970. The forgotten tone: a reply. *Le Maître Phonétique* 134: 29-31.
- Fromkin, Victoria (ed.). 1978. *Tone: a linguistic survey*. New York: Academic Press.
- Fourcin, A. J., and Evelyn Abberton. 1971. First applications of a new laryngograph. *Medical and Biological Illustration* 21: 172-182.
- Fry, D. B. 1979. *The physics of speech*. Cambridge: Cambridge University Press.
- Gandour, Jackson T. 1978. The perception of tone. In Fromkin 1978: 41-76.
- Gibbon, Dafydd. 1976. *Perspectives of intonation analysis*. Bern: Lang.
- Gimson, A. C. 1970. *An introduction to the pronunciation of English*. 2nd edn. London: Edward Arnold.
- Goldsmith, John A. 1976. *Autosegmental phonology*. Bloomington: Indiana U. L. C. (Published 1979 by Garland, New York).
- Greenbaum, Sidney, Geoffrey Leech, and Jan Svartvik (eds). 1980. *Studies in*

- English linguistics for Randolph Quirk*. London: Longman.
- Haggo, Douglas. 1983. Pitch-prominence and utterance function in stock auction speech. Paper read to the 5th Conference of the Linguistic Society of New Zealand, University of Auckland, May 1983.
- Haggo, Douglas, and Koenraad Kuiper. 1983. Review of F. Coulmas (ed.), *Conversational Routine*. *Linguistics* 21: 531-551.
- Haggo, Douglas, and Koenraad Kuiper. 1985. Stock Auction Speech in Canada and New Zealand. In Berry and Acheson 1985: 167-175.
- Halle, Morris. 1959. *The sound pattern of Russian*. The Hague: Mouton.
- Halle, Morris, and Jean-Roger Vergnaud. 1980. Three-dimensional phonology. *Journal of Linguistic Research* 1.1: 83-105.
- Halliday, M. A. K. 1967. *Intonation and grammar in British English*. The Hague: Mouton.
- Halliday, M. A. K. 1970. *A course in spoken English: intonation*. London: Oxford University Press.
- Hombert, Jean-Marie. 1978. Consonant types, vowel quality, and tone. In Fromkin 1978: 77-111.
- Hyman, Larry M. 1975. *Phonology: theory and analysis*. New York: Holt, Rinehart and Winston.
- Jackendoff, Ray S. 1975. Morphological and semantic regularities in the lexicon. *Language* 51:639-671.
- Jassem, W. and D. Gibbon. 1980. Re-defining English accent and stress. *Journal of the International Phonetic Association* 10: 2-16.
- Jespersen, O. 1928. *A modern English grammar on historical principles*, Part I: *Sounds and spellings*. London: Allen and Unwin.
- Kuiper, Koenraad. 1982. Once more with feeling: modifier repetition as a stylistic rule. *Linguistics* 20: 493-517.
- Kuiper, Koenraad. 1984. The pragmatics of stretching. Unpublished paper, University of Canterbury.
- Kuiper, Koenraad, and Douglas Haggo. 1984. Livestock auctions, oral poetry, and ordinary language. *Language in Society* 13: 205-234.
- Ladd, D. Robert, Jr. 1978. Stylized Intonation. *Language* 54: 517-539.
- Ladd, D. Robert, Jr. 1980. *The structure of intonational meaning*.

Bloomington: Indiana University Press.

Ladefoged, Peter. 1975. *A course in phonetics*. New York: Harcourt Brace Jovanovich.

Leben, William R. 1976. The tones in English intonation. *Linguistic Analysis* 2: 69-107.

Leben, William R. 1978. The representation of tone. In Fromkin 1978: 177-219.

Lewis, Jack Windsor. 1970. The tonal system of remote speech. *Le Maître Phonétique* 134: 31-36.

Liberman, Mark Y. 1978. *The intonational system of English*. Bloomington: Indiana U. L. C. (1975 Ph.D. thesis, MIT; published 1979 by Garland, New York.)

Liberman, Mark, and Alan Prince. 1977. On stress and linguistic rhythm. *Linguistic Inquiry* 8: 249-336.

Liberman, Mark, and Ivan Sag. 1974. Prosodic form and discourse function. *Papers from the Tenth Regional Meeting, Chicago Linguistic Society*, 416-427.

Maddieson, I. 1977. Universals of tone. *Stanford University Working Papers in Linguistic Universals*. (Cited in Anderson 1978.)

McCarthy, John J. 1981. A prosodic theory of nonconcatenative morphology. *Linguistic Inquiry* 12: 373-418.

Meyer-Eppler, Werner. 1957. Realization of prosodic features in whispered speech. *Journal of the Acoustical Society of America* 29: 104-106.
Reprinted in Bolinger 1972c: 385-390.

Miller, Jennifer, and Paul Tench. 1982. Aspects of Hausa intonation, 2: Continuous text. *Journal of the International Phonetic Association* 12: 78-93.

O'Connor, J. D. 1970. Review of Crystal 1969a. *Le Maître Phonétique* 133: 13-16.

O'Connor, J. D., and G. F. Arnold. 1973. *Intonation of colloquial English*. 2nd edn. London: Longman.

Pierrehumbert, Janet. 1980. *The phonology and phonetics of English intonation*. Ph.D. thesis, MIT.

Pike, Kenneth L. 1945. *The intonation of American English*. Ann Arbor: University of Michigan Press.

Pike, Kenneth L. 1948. *Tone languages*. Ann Arbor: U. of Michigan Press.

- Quirk, Randolph, and Sidney Greenbaum. 1973. *A university grammar of English*. London: Longman.
- Rochemont, Michael S. 1978. *A theory of stylistic rules in English*. Unpublished Ph.D. thesis, University of Massachusetts.
- Rochemont, Michael S. 1986. *Focus in generative grammar*. Amsterdam: Benjamins.
- Sag, Ivan, and Mark Liberman. 1975. The intonational disambiguation of indirect speech acts. *Papers from the Eleventh Regional Meeting, Chicago Linguistic Society*, 487-497.
- Schubiger, Maria. 1965. English intonation and German modal particles-- a comparative study. *Phonetica* 12: 65-84. Reprinted in Bolinger 1972c: 175-193.
- Schubiger, Maria. 1980. English intonation and German modal particles II-- a comparative study. In Waugh and van Schooneveld 1980: 279-298.
- Schuh, R. G. 1978. Tone rules. In Fromkin 1978: 221-256.
- Selkirk, Elisabeth. 1984. *Phonology and syntax: the relation between sound and structure*. Cambridge, Mass: MIT Press.
- Sommerstein, Alan H. 1977. *Modern Phonology*. London: Edward Arnold.
- Stevens, S. S., and J. Volkmann. 1940. The relation of pitch to frequency: a revised scale. *American Journal of Psychology* 53: 329-353.
- Trager, George L. 1964. The intonation system of American English. In Abercrombie *et al.* 1964: 266-270. Reprinted in Bolinger 1972: 83-86.
- Trager, George L., and Henry Lee Smith, Jr. 1951. *An Outline of English Structure*. Washington: American Council of Learned Societies.
- Vanderslice, Ralph, and Peter Ladefoged. 1972. Binary suprasegmental features and transformational word-accentuation rules. *Language* 48: 819-838.
- Waugh, Linda R., and C. H. van Schooneveld (eds). 1980. *The melody of language*. Baltimore: University Park Press.
- Wells, Rulon S. 1945. The pitch phonemes of English. *Language* 21: 27-39.